delete 删除

大数据取舍之道

The Virtue of Forgetting in the Digital Age



【英】维克托·迈尔-舍恩伯格◎著　袁杰◎译

VIKTOR MAYER-SCHÖNBERGER



浙江人民出版社

delete 删除

大数据取舍之道

The Virtue of Forgetting
in the Digital Age



【英】维克托·迈尔-舍恩伯格◎著　袁杰◎译
VIKTOR MAYER-SCHÖNBERGER

浙江人民出版社
ZHEJIANG PEOPLE'S PUBLISHING HOUSE

版权信息

本书纸版由浙江人民出版社于2012年12月出版

作者授权湛庐文化（Cheers Publishing）作中国大陆（地区）电子版发行（限简体中文）

书名：删除：大数据取舍之道

著者：（英）维克托·迈尔-舍恩伯格

字数：180000

电子书定价：24.99美元



《经济学人》说，在大数据领域，他是最受人尊敬的权威发言人之一；

《科学》说，若要发起一场关于这个问题的深入讨论，没有比他更好的发起者了。

他是欧盟互联网官方政策背后的重要制定者与参与者；

他是最早洞见大数据时代发展趋势的数据科学家之一；

他，就是 **维克托·迈尔－舍恩伯格**。

孜孜不倦的数据科学家



维克托·迈尔－舍恩伯格二十多年来一直致力于网络经济、信息与创新、信息监管、网络规范与战略管理的研究。
从维也纳大学到哈佛大学，从新加坡国立大学到牛津大学，世界上最著名的互联网研究学府都留下了他的足迹。

他是哈佛大学肯尼迪学院信息监管科研项目负责人，哈佛国家电子商务研究中心网络监管项目负责人。在哈佛大学任教10年后，2008年，他远渡重洋，来到新加坡国立大学担任信息政策研究中心主任，以崭新的视角洞察亚洲信息政策。期间还担任耶鲁大学、弗吉尼亚大学、圣地亚哥大学、芝加哥大学、维也纳大学等多所知名学府的客座教授。

而现在，在数据信息界孜孜不倦求索的维克托，是世界著名学府牛津大学网络学院互联网研究所治理与监管专业教授，牛津大学克伯学院教授研究员。法学、信息学与网络等多学科的跨界融合，让他不仅拥有严谨的思维，更拥有广博的视野。他先后有100多篇论文公开发表在《科学》《自然》等著名学术期刊上，同时也是多家出版机构的特约评论员，包括哈佛大学出版社、麻省理工学院出版社、通信政策期刊、美国社会学期刊等。

开大数据系统研究之先河



他说，世界的本质就是数据，大数据将开启一次重大的时代转型；
他说，大数据发展的核心动力来源于人类测量、记录和分析世界的渴望。
他说，从因果关系到相关关系的思维变革才是大数据的关键，建立在相关关系分析法基础上的预测才是大数据的核心。

当大数据还在众说纷纭时，维克托早已进行了系统而深入的研究。早在2010年，他已经在《经济学人》上和数据编辑肯尼思·库克耶一起，发表了长达14页的大数据专题文章，成为最早洞见大数据时代发展趋势的数据科学家之一。

从大数据时代的思维变革到商业变革，从大数据时代的隐忧到管理变革，他在大数据的蓝色海洋中预见了正在发生的未来。

大数据取舍之道



维克托的继父去世时留下了 16 000 张收藏的照片，这些照片都是他几十年来周游世界时的影像记录。

为了确定哪些照片需要保留，维克托制定了两条规则：首先，照片上有认识或可能认识的人；其次，照片拍摄得很漂亮。最终，他只留下了 53 张照片……

也许正是这一人生的片段，让他开始重新思考"被遗忘的权利"与互联时代的取舍之道，并最终成就了他的《删除》一书："过去正像刺青一样被刻在我们的数字皮肤上，遗忘已经变成了例外，而记忆却成了常态……"大数据时代，人类又该如何构建积极而安全的未来？

《删除》一经出版，即获得美国政治科学协会颁发的"唐·K·普赖斯奖"以及媒介环境学会的最高荣誉"马歇尔·麦克卢汉奖"，同时受到《连线》《自然》《华尔街日报》《纽约时报》等各大权威媒体的广泛好评。

大数据商业应用的引路人



维克托是真正的实战派，早在上大学期间，他就先后创立了两家数据安全和制作反病毒软件的公司，并担任总裁兼CEO。

由他的公司开发的反病毒程序，一经推出即一跃成为奥地利当时最畅销的软件产品。1991 年，他跻身奥地利软件企业家前 5 名之列，2000 年成为奥地利萨尔斯堡州的年度人物。

开阔的学术视野与系统的学术造诣，更让他不断为企业与商业应用提供强大的理论支持。他的咨询客户包括微软、惠普、IBM、亚马逊、facebook、twitter、VISA等大数据先锋们，所以在《大数据时代》一书中，他才能掌握最前沿、最崭新的大数

据应用案例，并对大数据的价值链与角色定位给予清晰的预见。

欧盟专家与真正的“亚洲通”



他是欧盟互联网官方政策背后真正的制定者与参与者；他是世界经济论坛、马歇尔计划基金会等重要机构的策略顾问；他是众多国家政府高层的信息政策智囊，包括俄罗斯商务部高层、瑞士联邦政府、荷兰政府高层等。

尤为重要的是，他特别熟悉亚洲信息产业的发展与战略布局，一直专注于信息安全、信息政策与战略的研究，先后担任新加坡商务部高层、文莱国防部高层、科威特商务部高层、迪拜及中东政府高层的咨询顾问。

正如大数据提供的不是最终答案，只是参考答案一样，大数据正在改变我们的生活以及理解世界的方式，正在成为新发明和新服务的源泉，而维克托·迈尔-舍恩伯格更多的大数据研究也在蓄势待发......

维克托·迈尔-舍恩伯格 作品系列

大数据时代



删 除



湛庐文化 CheersPublishing 特别制作

中文版序 大数据取舍之道

几乎没人能够想象，没有了互联网的世界将会怎样。互联网赋予了公民权利，大大提升了经济效率，带来了不计其数的创新爆发。数以亿计的人正在使用互联网购物，便捷轻松地与人沟通，分享观点。

然而，互联网的全球化意味着，完整的数字化记忆以及“记忆成为常态”所带来的诸多挑战，不仅存在于伦敦和旧金山，在北京和上海也同样令人关注。跨越了地理上的界线，完整的数字化记忆正在挑战着我们所有人。因此，本书所传达的信息对于亚洲的互联网用户同样也很重要。

有人认为，完整的数字化记忆在非民主社会更加危险，因为非民主社会中公民无法直接限制政府利用从网上收集到的海量私人信息。但是，这种观点却遗漏了关键的一点，那就是：大量数字化的私人信息不仅可能在今天被滥用，在几年甚至几十年后仍然可能被滥用。所以，从某种程度上讲，一个人是否生活在民主社会，与数字化记忆是否被滥用并无必然关系。如果私人信息被抓取后，以数字形式在未来被储存了很久很久，那么，即便生活在民主社会，它仍然像一颗定时炸弹——因为，未来任何非民主政府都可以利用这极具价值的信息权力宝藏。

假设真正的危险并非来自于政府机构滥用信息权力，而是来自于巨大的商业机构——这些商业机构借助自己位于全球信息流中战略瓶颈的优势地位，获得我们相当多的私人信息流，那么又将会怎样呢？就像对政府那样，我们只能希望，像Google和facebook这样的巨型网络公司能够清楚地知道，他们对于收集并保存的几十亿项私人信息所负有的责任。

我们不知道未来将会怎样：会拥有什么样的商业平台以及他们会以何种目的去挖掘自己积累的私人信息宝藏？他们会不会滥用我们的私人信息？正因如此，我们要开始思考减少我们的数字足迹：不是通过戒掉互联网，而是通过塑造互联网及其服务，以使得数字信息能够真正在一段时间之后被渐渐遗忘。如此看来，数字化记忆给所有互联网用户带来了一连串的挑战，不管他们生活在世界的哪个角落。

这也许是我们需要牢牢记住的最根本最重要的信息。我不是新勒德分子[1]，不鼓吹与全球互联网断开连接。我的建议恰恰相反，那就是：充分运用互联网使其发挥全部潜能；并且，通过给互联网设置遗忘的功能，以确保互联网被那些大权在握之人当作工具滥用的可能性越来越小，而让互联网一直充当创新、沟通和赋予权利的工具。

我在书中描述过的许多情况每天仍然在发生：被互联网抓取到的信

息量在突飞猛进地增加。在2000年，世界范围内绝大部分信息是以模拟形式记录的，而今天，世界上90%以上的信息是数字形式的——因此能够毫不费力地进行存储、加工、操作和发送。

不仅在电脑上浏览互联网时如此，智能手机里的多种传感器也是如此。它们能够抓取日期、时间、地理位置，甚至是从汽车的加速度、行驶方向到大气压强等一切信息。

手机上安装的应用程序使得置于床垫下面的智能手机能够获取我们的夜间睡眠模式，而嵌在小手镯里的芯片能够记录我们的运动。

一家美国运动服装公司现在甚至在出售拥有内置显示器的滑雪眼镜，这款眼镜不仅能够捕捉我们的位置以及我们滑下坡时的速度，而且还能告诉我们要找的朋友在哪里。

这一切都很重要，运用强大的统计工具，我们能够理解抓取到的大量数据，而且能够提取其中一些潜在的价值。“大数据”运动就像互联网一样，将会在世界和社会的运行方式上带来跨越式的变革。借助大数据，我们会有更好的医疗措施，更少的意外事故，更高效的市场格局，以及对社会更好的理解与认知。但是，“大数据”也会增加抓取更多数据以及将它存储更长时间的压力，对数据的过分痴迷会让人唯恐丢失一些可能获得的潜在价值。

在这个背景下，完整数字化记忆的隐患需要的不仅仅是重复提醒——它们需要被不断强调！我相信，《删除》将会在未来很多年对我们理解互联网具有深远的价值。只有这样做，才能确保互联网一直是经济和社会进步的智慧源泉，而非大权在握之人赋权的工具。而且，只有这样做，才能确保在进入“大数据”时代后，我们不必害怕会失去隐私，会失去人性化的生活与行动。

我是一个乐观主义者，所以我会乐观地号召所有人，要在大数据时代始终记得遗忘的美德。

---

[1] 勒德，19世纪初参加捣毁机器运动的英国手工业工人，强烈反对机械化或自动化。后用勒德主义者指仇视一切新奇发明乃至科技进步的人。——作者注

推荐序 因意义而智慧

中国社科院信息化研究中心秘书长
《互联网周刊》主编
姜奇平

维克托的《删除：大数据取舍之道》到底在讲什么？

英美人写书，总爱从细节到细节，进行无穷无尽的经验归纳，就像素描一样，一笔一笔地在相近的地方重复描摹。大家的时间都有限，让我用演绎的方法，谈谈这本书的逻辑，就像速写那样，几笔让你看出画的是什么。

大艺术家罗丹说："什么是雕塑？就是在石料上去掉那些不要的东西。"删除，就是当大数据这堆石料越来越多后，去掉那些不要的东西。去掉不要的，为的是让雕像留下来。雕像就是意义所在。简单地说，大数据的取舍之道，就是把有意义的留下来，把无意义的去掉。

这本书到此就可以算读完了。

不过也许你还有五分钟时间，那就可以听我多说两句，以便对这本书更加知其然。

《删除》这本书为什么值得你读？一个重要原因是它来得恰是时机。大数据刚刚兴起，在你被数据灌满之前，先教你如何删除，省得你的硬盘和大脑，在数据海啸到来后，成了无用信息的垃圾场。

《删除》这本书的缘起，是2007年作者偶然产生的"被遗忘的权利"这个灵感。作者的一位朋友向他鼓吹计算"无处不在"，他却反其道强调遗忘的重要，也就是不要让数据"无处不在"。

大家都有这种体会，信息少的时候，脑子还清楚；等到信息爆炸后，脑子反而乱了。《删除》这本书，像念经似的，翻来覆去就在重复两个词：记忆，遗忘。这不奇怪，因为记忆与遗忘的斗争，构成了数据的核心矛盾。正如作者所说："在数字时代，或许人类发生的最根本的改变，就是记忆和遗忘的平衡已经反转了。将信息提交给数字存储器已经成为默认状态，而遗忘则成了例外。"以前数据过少的时候，想记、该记的没记下来；如今数据过多的时候，想忘、该忘的忘不了。大脑就像仓库一样，没用的东西占的地方太多了，需要清一清。《删除》就在教你如何清扫大脑内的垃圾。

作者写这本书，是为了向你说明"怎样才能够确保在数字时代的未来我们还会记得遗忘的重要性"，并为此"恢复我们遗忘的能力"。也就是不要让乱七八糟的东西过多占据我们的大脑和硬盘。作者教给你的，相当于罗丹的石料去除术。

如果你还有十分钟，我还想跟你谈谈，如何理解这本书的所以然。

如果你光就这本书读这本书，可能也会被绕进去。因为作者把主要篇幅，都用在教你如何去掉石料中不需要的部分，但是并没有讲透石料中保留的雕像到底是什么。其实，你只有明白自己到底要雕的是什么，才能明白去掉石料的所以然。换句话说，你只有理解了在大数据中，你需要的是什么，以及如何判断这种需要，才能举一反三地明白到底为什么要去掉那些不需要的。否则，你就会一直凿，直到石料都快没了的时候，才忽然问自己：咦，我凿它干什么呢？

雕像是意义。意义是决定哪些数据值得保留（记忆），哪些数据需要删除（遗忘）的根据。罗丹在雕塑的时候，不是在做刀削面，不是为了把东西全去掉完事，而是为了一件正事，就是雕出像来。

维克托在新作《大数据时代》中解决了前面遗留的问题，他对大数据进行了详细的解释。大数据只是客体，本身并不能决定自己有用还是没用。有用还是没用，是相对于主体来说的。对人有意义的数据，就是雕像，就是该保存的回忆；对人没意义的数据，就是应去掉的石料，就是该删除的垃圾。对这个人有意义的，对那个人可能没有意义。因此，不同的人，要根据自己认同的特殊意义，来决定数据的取舍。智慧与非智慧，是取舍的结果。取有意义的，舍无意义的，叫智慧；取无意义的，舍有意义的，叫愚蠢。大数据因有意义而变得智慧，这就是取舍中“取”这一半的道，它是《删除》“舍”那一半道的所以然。合在一起，才是取舍之道。

进一步，取之有道，这个道可以点破了说，就是要明白在决定遗忘什么的时候，先明白到底要回忆什么。其中的道，在于柏拉图意义上的回忆。其实，回忆与遗忘，是欧洲文化中一个经典隐喻，有特别的含义。例如贝尔纳·斯蒂格勒《技术与时间：爱比米修斯的过失》谈的记忆与遗忘，就是这种意义上的。对柏拉图来说，真正的回忆，不是对现世的回忆（即一件事一件事的回忆），而是对彼岸的回忆（即对“人是什么”的认同，也就是对意义的认同）。他在《斐德若》中说“见到尘世的美，就回忆起上界里真正的美”。还原到现实中来，就是各种回忆中与自己的存在意义最息息相关的回忆。在数据这堆石料背后，真正存在的“雕像”，就是人的意义本身。

罗丹心目中的雕像，反映的正是他所认同的意义。他根据这种认同，来决定“删除”哪些石料，那一定是同意义无关的。同样，在未来的大数据时代，每个人都是罗丹，每个人只需要保留那些可供“回忆”有意义事物的数据，而删除那些无意义的数据。删除的对与不对，有效与无效，完全取决于我们对自己的认识。从这个意义上说，“认识你自己”这

句阿波罗神庙上的箴言，应该成为删除行为的真正标准。

删除的最高境界，就是不删除，也就是知道你需要保留的东西。

目录

第一部分 大数据时代为什么要进行信息的取舍

01 当遗忘变成例外，而记忆成了常态：大数据时代的隐忧

对于人类而言，遗忘一直是常态，而记忆才是例外。然而，由于数字技术与全球网络的发展，这种平衡已经被打破了。如今，往事正像刺青一样刻在我们的数字皮肤上，遗忘已经变成了例外，而记忆却成了常态……

在这个记忆已经成为常态的时代，难道每个公开自己信息的人只能永远对信息束手无策吗？我们真的想要一个由于无法遗忘，而永远不懂得宽恕的未来吗？

未来可能遭遇到的悲剧，会对我们现在的行为产生寒蝉效应。借助数字化记忆，圆形监狱能够随时随地监视我们。

事实上，并不是技术迫使我们去记忆。技术只是促进了遗忘的终止——而只有当我们迫切想要记住的时候，遗忘才会被终止。

喝醉的海盗

史黛西·施奈德（Stacy Snyder）最大的梦想就是成为一名教师。在2006年春天，这位25岁的单身母亲已经完成了她的学业，并且对她未来的事业充满了期待。但是她的梦想破碎了。她心仪的学校明确地告诉她，她不能成为一名教师，虽然她已经修满了所有的学分，通过了所有的考试，完成了所有的实习训练——而且在许多方面成绩优异。然而，她被取消了当教师的资格，理由是，她的行为与一名教师不相称。

她的行为？校方指的是一张网上的照片，照片里的她头戴一顶海盗帽子，举着塑料杯轻轻啜饮着。史黛西·施奈德曾将这张照片放在MySpace她的个人网页上，并且取名为“喝醉的海盗”。这张照片她是给朋友看的，也许只是为了搞怪而已。然而，在史黛西实习的那所大学里，一位过度热心的教师发现了这张照片，并上报给校方，校方认为网上的这张照片是不符合教师这个职业的，因为学生可能会因看到教师喝酒的照片而受到不良影响。于是，史黛西打算将这张照片从她的个人网页上删除。但是危害已经发生了：她的个人网页已经被搜索引擎编录了，而且，她的照片已经被网络爬虫（web crawler）程序存档了。互联网记住了史黛西想要忘记的东西。

后来，史黛西控告了这所大学，但也没能最终胜诉。她认为，将那张照片放在网上并不能说明一名未来教师的不称职和不专业。毕竟，在那张照片中并没有显示塑料杯里装的究竟是什么东西。即便显示了，史

黛西，这位拥有两个孩子的单身母亲也早已达到了在私人聚会上喝酒的年龄。

然而，这个案例与大学是否应该给史黛西教师资格的正确性或愚蠢程度无关，而是与某种更为重要的东西有关，那就是遗忘的重要性。

一个没有遗忘的时代

有史以来，对于我们人类而言，遗忘一直是常态，而记忆才是例外。然而，由于数字技术与全球网络的发展，这种平衡已经被改变了。今天，在广泛流行的技术的帮助下，遗忘已经变成了例外，而记忆却成了常态。这种局面是如何发生的，为什么会发生，对于我们个人和我们的社会而言这一局面的潜在后果是什么，以及针对这一局面，如果可能的话，我们能够做些什么，这些都是本书关注的问题。

对于某些人而言，史黛西·施奈德的遭遇可能听起来只是个例外，但其实不然。自此之后发生了许多起非常窘迫的案例，甚至是法律诉讼——从无法让互联网忘记十多年前细微证据的知名大律师，到由于在facebook上抱怨工作无聊而因此失业的英国小姑娘。

截至2008年，在MySpace上拥有个人网页的人数超过1.1亿，他们都和史黛西·施奈德一样。而且MySpace并不是网民唯一青睐的社交网站。MySpace的直接竞争者facebook，截至2009年年初，已经有1.75亿个个人用户创建的在线网页。尽管facebook与MySpace都主要锁定美国市场（虽然这种局面正在改变），但是这种现象并非仅仅发生在美国。Google公司旗下的社交网站Orkut，已经拥有1亿多名用户，而且大部分用户在巴西与印度。世界各地的其他多个网站还拥有至少2亿名用户。这些数字反映了一个更为普遍的趋势。

互联网浪潮的第一个阶段早已随网络泡沫及其破灭而告终，在这个阶段中，互联网的发展与访问信息以及通过全球网络与他人互动息息相关（你可以称其为Web 1.0时代）。而到2001年，用户开始意识到互联网并不仅仅只是一个接收信息的网络，而且还是你可以生成信息并与你的同龄人共享信息的网络（通常被称为Web 2.0时代）。年轻人尤其喜欢Web 2.0的这些性能。

到2007年年底之前，美国的一家趋势调查机构皮尤研究中心（Pew Research）发现，三分之二的青少年“参与了互联网上一个或者多个范围广泛的内容创建活动”，而且在创建和共享内容上，女孩要多于男孩。每天，facebook平均每秒钟都会收到来自世界各地用户的1000万个网络应用请求。约翰·帕尔弗雷（John Palfry）与乌斯·加瑟教授（Urs Gasser）曾充分论述，公开个人的信息都已经深深地嵌入了全球的青年

文化中——无论是登入facebook的次数、个人日志与评论、照片、好友状态，还是情侣关系（“恋爱中”或者“只是朋友”）、内容偏好与认同（在线照片或者“标签”）、个人的地理位置（通过“地理标签”或者类似Dopplr的网站可获知），或者仅仅只是微博更新（比如twitter）。

随着这些年轻人逐渐长大，以及越来越多的成年人拥有类似地特征，史黛西·施奈德的遭遇将会不断被重演，而且不仅对于整整一代人是如此，对于我们整个社会也将如此。

抹不掉的致幻剂阴影

虽然Web 2.0刺激了这一局面的发展，但是传统出版业在互联网力量的推动下也导致了惊人相似的结果。让我们来看看安德鲁·费尔德玛（Andrew Feldmar）的遭遇。

这位六十多岁的加拿大心理咨询师生活在温哥华。2006年的某一天，一位朋友从西雅图国际机场赶过来，费尔德玛打算穿过美国与加拿大的边境去接他，就像曾经上百次做过的那样。但是这一次，边境卫兵用互联网搜索引擎查询了一下费尔德玛。搜索结果中显示出了一篇费尔德玛在2001年为一本交叉学科杂志所写的文章，在文中他提到自己在20世纪60年代曾服用过致幻剂LSD。因此，费尔德玛被扣留了4个小时，其间被采了指纹，之后还签署了一份声明，内容是他在大约40年前曾服用过致幻剂，而且不准再进入美国境内。

安德鲁·费尔德玛是一位没有犯罪记录、拥有学识的专业人员，他知道当年服用致幻剂确实违反了法律，但是他坚称自1974年以来就一直没再服用过。当边境卫兵拦下他时，这件事已经过去快40年了。对于费尔德玛而言，那是他生命中一段早已远去的时光，一个他认为已被社会遗忘了许久、与现在的他完全不相干的过错。但是，数字技术已经让社会丧失了遗忘的能力，取而代之的则是完善的记忆。

有人会说，史黛西·施奈德的痛苦主要是由她自己造成的。她将她的照片放在自己的网页上，并且加了一个引发歧义的标题。或许，她并没有意识到全世界都能找到她的网页，而且竟然在她删除照片很久之后，仍有可能通过互联网档案访问她的照片。作为互联网时代的一员，她也许应该更谨慎地考虑一下，哪些内容可以在互联网上公开。不过，安德鲁·费尔德玛的遭遇却与她不同。年近七十的他可不是十几岁的互联网发烧友，而且可能从来没预料到，他在那样一本晦涩杂志上发表的文章，居然能在全球化的网络上如此容易地被找到。对他而言，成为数字化记忆的受害者完全是一个可怕的突然袭击。

但是，即便史黛西与安德鲁能预见这个结果，难道这就意味着每个

公开自己信息的人只能永远对信息束手无策吗？关于互联网是否会遗忘以及何时遗忘，难道我们没有一点发言权吗？我们真的想要一个由于无法遗忘，而永远不懂得宽恕的未来吗？

“现在，青春期时一个愚蠢的错误会造成严重的后果，而且这个不良记录的阴影会一直笼罩着他们的余生。”家庭教师协会（PTA）的联合主席凯瑟琳·戴维斯（Catherine Davis）评论说。如果我们不得不担心，关于我们的所有信息，将会被保留得比我们的寿命更长，我们还会对琐碎的流言蜚语发表我们的观点吗？我们还会轻松地分享个人的经历吗？我们还敢做出各种政治评论吗？我们还会进行自我审查吗？完善记忆的寒蝉效应[1]会改变我们的行为。

史黛西与费尔德玛表示，今后将会以完全不同的方式谨慎行事。“当心你发布在网上的信息。”史黛西说。而费尔德玛则更富见地地感叹，“我要警告所有人，你留在网络上的电子足迹将在某一天对你造成伤害。那是无法被擦除的。”

取舍之道

不再遗忘的后果，比人们建立并维护多年的声誉遭受一次正面攻击的后果要深远得多，也更让人苦恼。如果我们过去所有的行为，无论是否违法，都一直存在，那么当我们进行思考与决策时，怎样才能从自己的过去中解脱出来呢？完善的记忆会使我们对待他人像对待自己那样不宽恕吗？

尽管如此，史黛西与费尔德玛仍然自愿地公开关于他们遭遇的信息。从严格意义上来讲，他们也需要为这种公开的后果承担责任。然而，我们往往在公开自己的信息时，并没有真正意识到我们正在“公开”自己的信息。

Google记得你的一切

在德国城市爱森纳赫（Eisenach）城外坐落着一家名叫MAD的巨型迪斯科舞厅，它能容纳4000名顾客。当顾客进入这家舞厅时，他们必须出示他们的护照或者由政府发布的身份证；这些顾客的详细资料都被输入到一个数据库中，连同一起被存入的还包括他们的数字脸部照片。每名顾客都会有一张专用的支付卡，他们只能用这张卡支付在MAD餐厅及多个酒吧中消费酒水与食品的账单。每次交易都会被添加到顾客的永久数字记录中。据一次电视报导，2007年年底之前，MAD数据库存储的信息包括一万三千多名顾客与几百万次交易。60台数字摄像机持续地拍摄迪斯科舞厅及周围的每个角落；记录与存储的录像所占的硬盘空间超过8000G。顾客的实时信息、他们的交易行为，以及他们的消费偏好

都被显示在一个专用控制室里的大屏幕上，这个控制室中的情景仿佛出自007电影一样。管理人员很自豪地说，每周7天每天24个小时，当地的警察都能通过互联网在线访问MAD舞厅硬盘里存储的信息。

即便有，也只是少数顾客才能意识到他们的一举一动都被记录了，并且会被保存多年，同时第三方机构还可以访问这些信息——这给成千上万名不知情的顾客造成了一种深远的信息阴影。

再举一个更为普遍的例子，互联网搜索引擎。Google、Yahoo！、微软搜索、Ask.com以及对万维网（World Wide Web）进行编录的其他许多搜索引擎，通过对网页的逐个搜索，使得所有人仅在搜索框中键入一两个关键词就能访问这些网页。我们知道，同时假设搜索引擎也知道大量的信息，而且这些信息可以通过全球互联网上的网页进行获取。多年来，这种易于使用且功能强大的搜索引擎，成功地将世界各地的信息资源展现在几十亿用户面前。然而，搜索引擎所记住的，要比网页上所发布的信息多得多。

2007年春天，Google承认，直到那天他们一直在存储每位用户曾经键入的每次搜索请求，同样被存储的还有每位用户随后点击访问的每一条搜索结果。

通过保存大量被整齐排列的搜索关键词（Google每月收到的搜索请求大概有300亿次），Google能够将这些数据与人口统计资料联系起来。比如，Google能够显示出搜索请求的趋势，甚至是几年以后的趋势。Google能告诉我们，2006年秋天在美国印第安纳州首府印第安纳波利斯人们多久会搜索一次“伊拉克”，或者2007年圣诞季在亚特兰大的中产阶级中搜索最多的关键词是什么。不过，更为重要的是，通过智能地整合登录数据、cookies数据与IP地址，Google能够以极高的准确率将时间跨度很远的多次搜索请求与某个人关联起来。

结果非常惊人。Google能够识别出我们每个人在搜索什么，什么时候搜索的，以及我们可能觉得足够准确并点击进去的搜索结果是什么。Google知道我们生活中的任何重大改变——在2000年你婚礼之后买了一套房子，在2003年有一次可怕的非典恐慌，在那一年年底你有了一个新出生的宝宝。但是，Google也知道关于我们的详细细节，哪怕是我们已经忘记了很久的细节，被我们大脑当作不相干的东西而丢弃掉的细节。但是，这些被遗忘的细节却清楚地显现出了我们的过去：可能当考虑对前雇主采取法律手段时，我们曾经搜索过某位职业律师、研究过某个心理健康问题、找过某本最新的小说；或者当与一个人保持恋爱关系时，我们却预订过某个隐蔽的汽车旅馆与另一个人约会。这些“信息位”中的每一条都已经被我们抛于脑后了，但是Google并没有忘记。可以毫不夸

张地说，**Google**对我们的了解比我们自己能够记住的还要多。

Google已经宣布，他们将不再永久地保存用户的记录，而是在保存9个月之后将记录匿名化，因此便模糊了他们完整记忆中的一部分信息。将用户搜索记录保存9个月，仍然为Google提供了一种非常有价值的信息资源，他们会在认为适当的时候使用，而且一旦过了保存期，Google的承诺也仅仅是模糊搜索请求中关于用户的标识符，并非清除实际的搜索请求本身，也不是清除存储它时的背景信息。所以，在Google无法告诉我五年前我搜索的关键词是什么，以及我点击的搜索结果是什么的时候，他们可能仍然能够告诉我一个相对较小的群体，在5年前的4月10日晚上都搜索了什么。这个群体可能是一个与我收入水平相当，并且在我家附近拥有一所房子的中年男性群体。依据群体的规模，这些数据仍然能够透露我作为一个个体的很多信息。而且与史黛西·施奈德与安德鲁·费尔德玛相比，我们很少有人知道，Google对我们的搜索保存了如此精确的记录。

Google并不是记录用户信息的唯一一家搜索引擎。雅虎大约每个月有100亿次搜索请求，同时也是世界上第二大互联网搜索提供商。据称，雅虎像微软的搜索引擎“必应”一样，保存了类似地个人搜索请求记录。

大数据的信息力量

有些机构能够接近完美地保存关于我们每个人如何使用他们服务的记忆，而且他们利用这种信息权力理所应当，搜索引擎便是这类机构中最强大的例子。当然，其他的机构也收集并且保存了大量关于我们的信息。

大型国际旅行预订系统同样记住了我们已经忘记很久的信息，这些系统正被在线旅游网站，比如美国艾派迪公司（Expedia）或者Orbitz网站，以及世界上成千上万家传统旅游社所利用。每个人通过这些机构办理的每次航班预订都被存储在他们的电脑中好几个月，即便我们并没有预订过航班，我们的旅游信息也会被记录。他们的记录能够在我们计划上次假期6个月之后，告诉我们上次考虑的目的地是哪里，选择了什么航班，或者我们想跟谁同行（即使那个人可能从来没有接触过这种服务，而且可能也不知道她成了这样一个人选）。他们记住了我们已经忘记了很久的信息。

信用机构存储了几亿美国公民的大量信息。美国最大的营销信息提供商为其数据库中的2.15亿人，提供了高达1000个数据点。我们还看到之前完全无关的数据源，现在被组合在一起。隐私专家丹尼尔·索洛夫

（Daniel Solove）描述了一家公司，这家公司通过来自全球2万个不同数据源的信息，能够提供一种关于某个个体的综合看法。而且，它会保留这些信息，即便个体会怀疑这种信息的准确性。医生会保存医疗记录，而且迫于经济与管理上的压力，他们不得不将几十年来高度私人化的信息置于数字存储器中，将其数字化。事实上，不是只有私营企业才致力于这种完善的数字化记忆，执法部门也存储了几千万人的生物识别信息，即便他们从来没有因犯罪而被指控，但大部分这些既敏感又可以搜索的记录从来没有被删除过。

并不只有美国在制造大大超出了我们所有人类心智容量的数字化记忆。在英国，有420万个摄像机在勘查公共场所，记录着我们的行动。到目前为止，存储容量与面孔识别能力上的极限制约了它的可获取性，但是新技术将很快被用于实时地识别个人身份。据英国广播公司报道，这里提到的技术据传是由拉斯维加斯的赌场开发出来的。

政策制定者不去想如何使公民免受傲慢监视与记忆的伤害，反而强制私营企业完善数据采集装置，以完善关于我们所有人的数字化记忆。同时，让从情报部门到执法部门的公共机构能够便捷地获取这些信息。

这可能还仅仅只是一个开始。已经有一些手机装载了全球定位系统（GPS）接收器，使得准确定位并跟踪我们的行动成为可能。许多公司正在销售GPS跟踪设备，这样担心子女的父母们就能够跟踪他们十几岁的孩子，多疑的夫妻用它来跟踪自己本不该被怀疑的伴侣。第一部带有GPS芯片的数码相机已经面世，能够将位置信息添加在我们拍摄的每张照片与每段视频上，所以不仅日期与时间，还有我们留下回忆的地点都会被印刻在数字化记忆中。在不久的将来，我们身边的东西都将会附带上微小且价格合理的传感器，这种传感器能够记录这些东西的下落。因此，很可能会出现一种第三方设备，它们不仅能够获得我们在何地的完整数字化记忆，还能够获得我们在何时，以及如何与我们身边的东西进行互动的数字化记忆。很有可能，一种比以往更为广泛的，关于我们行动的踪迹将会被收集起来，并被保存在数字化记忆中。

人类住进了数字圆形监狱

当然，这并不全是坏事。实际上，在许多方面，能够负担得起的、全面的数字化记忆，对我们每个人以及社会都是有益的。这种存储器能够帮助我们记录思想并捕捉某些时刻，当我们日后重温那些时刻时，将会为我们带来欢乐与满足。它抵制了人类健忘特性恼人的一面，提醒我们记住生日与结婚周年纪念日，誓言与承诺。

当有关公司的产品是如何被生产和销售的信息，不仅保存在雇员飞

逝的记忆中，而且被保存在一种更为持久的信息存储器中时，公司的运作会更加高效；这种持久的信息存储器——从常规的笔记本到数字信息库，能够同时被现在与未来的同事所共享。当生产商能够用客户过去的偏好来预测未来的需求时，市场可能变得更经济。最终，我们的社会也能够从这种记忆中获益，因为它帮助避免了再次犯下代价高昂且又危险的错误。正如许多人所说，从历史中学习需要社会拥有记忆的能力。

然而，当我们听到史黛西·施奈德与安德鲁·费尔德玛的遭遇时，我们会感到心神不安。当我们意识到Google的数字存储器，或者信用机构、旅行预订系统、电信运营商与执法部门的存储器有多么强大和完整时，我们被震惊了。同样重要的是，我们收集从个人电脑到数字视频录像机，从相机记忆卡到数码音乐播放器的一切数字信息，它们已经膨胀得像气球一样，使得我们能够获取我们已经抛之脑后的信息。我们感觉到的是——遗忘的终止，以及向记住这种常态根本性的转移。当记忆具有无可争议的好处时，事实上，其中很多的记忆可能会导致可怕的后果。

多年来，隐私专家一直在警告可能会出现一些后果。实际上，美国现代隐私争论的诞生，便起源于对完整数字化记忆的反对。亚瑟·米勒（Arthur Miller）于1971年出版的著名作品《隐私的侵犯》（The Assault on Privacy），正是由于美国联邦政府打算建立一个国家数据库计划而促成的产物。世界上第一部数据隐私法案，诞生于德国黑森州（Hessia），它的通过是对德国政府类似计划的直接回击。

其他人也对越来越多地使用监视技术跟踪人类活动的情况提出了义正言辞的批评。他们警告说，这有可能会导致英国哲学家杰里米·边沁（Jeremy Bentham）所谓的“圆形监狱”数字版。在这种监狱中，狱警能够在犯人不知道自己是否被监视的情况下监视犯人。边沁认为，这种监狱结构将能迫使犯人好好表现，而且这种方法使得社会付出的代价最小。因此，这是一种“新的监视模式，其信息权力之大前所未有”。社会学家米歇尔·福柯（Michel Foucault）采用了边沁的概念，并且表示，全景敞视建筑已经远远超越了监狱本身以及边沁关于实体监狱结构的观念，这种全景敞视建筑现在被更为抽象地用作在我们的社会中施展信息权力的工具。在这里，通信理论家奥斯卡·甘迪（Oscar Gandy）将圆形监狱，与我们时代中日益明显的、向大规模监视发展的趋势联系在一起。这种圆形监狱塑造了我们现在的行为：我像被人监视时一样行动，即便并没有人监视我。

完整的数字化记忆代表了一种更为严酷的数字圆形监狱。由于我们所说与所做的许多事情都被存储在数字化记忆中，并且可以通过存储器进行访问，因此，我们的言行可能不仅会被我们同时代的人们所评判，

而且还会受到所有未来人的评判。史黛西与安德鲁的惨痛经历，Google与其他搜索引擎掌握的大量与我们有关的数字化记忆，这些事实的刺激让我们变得极度警惕——换言之，未来可能遭遇到的悲剧会对我们现在的行为产生寒蝉效应。通过数字化记忆，圆形监狱能够随时随地监视我们。

取舍之道

毫无疑问，个人隐私被侵蚀是我们这个时代正面临的一种核心挑战，但这并不是本书关注的重点。本书的关注点与众不同，从狭义角度和广义角度来讲都是如此。

它关注的既不是大规模监视的危险，也不是关于对个人言行的跟踪，或者自我暴露的兴起；它关注的是遗忘与记忆在我们社会中扮演的角色，以及它们的角色如何发生变化；它还关注这种角色的变化可能会造成的潜在影响，以及对此我们能够做些什么，如果能够做的话应当怎样做。

让我们学会遗忘

遗忘在人类决策过程中扮演了重要的角色。遗忘使得我们能够及时地进行行动，知晓往事，但又不受往事的束缚。完善的数字化记忆，可能会让我们失去一项人类重要的能力——坚定地生活在当下的能力。博尔赫斯的短篇小说《博闻强识的富内斯》（*Funes the Memorious*）展现了这一论点。由于一次骑马的事故，年轻人富内斯失去了遗忘的能力。通过惊人的阅读，他积累了大量关于经典文学作品的记忆，但却无法超越字面的文字去领会作品的内涵。博尔赫斯暗示，一旦我们拥有了完善的记忆，我们将不能进行概括与抽象化，这会让我们一直迷失在过去的琐碎细节中。博尔赫斯的小说仅仅只是假想，我们不妨来看看现实中的真实情况。

在研究者最近发表的报告中，一位病人简称为AJ，她是美国加州一位41岁的妇女，天生就没有遗忘的能力。自她11岁开始，她几乎能记住每天发生的事情——她记住的不是过去一天的大致感觉，而是能够惊人地记得让她苦恼的详细细节。她清楚地记得，三十年前的一次早餐吃了什么；她能够回忆起谁在什么时候给她打了电话；她能够记得在20世纪80年代看过的电视节目每一段都演了什么。她甚至不需要努力地回忆就能想起这些。记忆对她而言很简单——她的记忆是“不可控且自动的”，就像一部“永远不会停止的”电影。

这带给AJ的并不是超常的能力，恰恰相反，她的记忆不断地限制了她做决定与前进的能力。她记住的信息包括自己经历的、感觉到的以及

想到的事情。那些在存储与回忆大量信息方面拥有超常能力的人，其实很想关闭他们记忆新事情的能力，至少是想暂时关闭。持续浮现的往事让他们感觉受到了束缚，这种束缚非常严重，以致于约束了他们的日常生活，限制了他们的决策能力，阻碍了他们与正常人建立紧密的联系。当这种影响由更为完整且更易获取的外部数字化记忆所引起时，影响可能会更强。如果回忆太清晰，即便这种回忆是为了帮助我们的决策，可能也会使我们困于记忆之中，无法让往事消逝。这反而会让我们像博尔赫斯笔下的富内斯一样，没有能力进行抽象的思考。这是记忆带来的意外诅咒。

取舍之道

遗忘不仅仅是一种个人的行为，我们这个社会也会遗忘。往往这种社会性的遗忘能给那些失败过的人第二次机会。如果原先的社会关系不能让人们感到幸福，我们可以让他们尝试建立新的社会关系。在商业领域中，随着时间流逝，破产会被逐渐遗忘。甚至在某些情况下，罪犯在经过足够长的时间之后，之前的犯罪记录也能从他们的档案中变得模糊。

通过很多这种模糊掉了外部记忆的社会遗忘机制，我们的社会能够接受随着时间不断发展的人们，因此我们才有能力从过去的经历中吸取教训，并调整我们的行为以融入未来的社会。

尽管遗忘对于人类非常重要，但是在数字时代，我们却正在经历一种重大的转变，从遗忘为常态转移到记住为常态，而且目前为止，这一现象得到的关注非常有限。回到1998年，J.D.拉西卡（J.D.Lasica）在在线杂志《沙龙》（*Salon*）上发表了一篇非比寻常的文章，题目是《互联网永远不会忘记》（*The Net Never Forgets*），并总结称“我们的过去正像刺青一样刻在我们的数字皮肤上”。最近，利亚姆·班农（Liam Bannon）、让-弗朗西斯·布兰切特（Jean-François Blanchette）与黛博拉·约翰逊（Deborah Johnson）已经开始揭露遗忘终止所带来的负面效应。在本书中，我通过考察当前最岌岌可危的现象，去评估、探索最有可能的补救方法；同时，在过去的人类历史与现在的数字时代中去探索记忆与遗忘。

取舍之道

本书无法穷尽遗忘的所有方面，包括它与个性、尊严、选择以及人类随着时间发展的能力之间的关系。当然，也无法提供一个终极解决方案。但是书中将提出一个我认为简单却有力的论点：这种从遗忘到记住的转变是意义重大的，而且如果不加以解决，可能会在未来同时给我们个人与整个社会造成严重的后果。

不过，这样的未来并不是不可避免的。技术并没有迫使我们去记忆，技术只是促进了遗忘的终止——而只有当我们迫切想要记住的时候，遗忘才会被终止。事实上，正是我们自己造成了遗忘的终止。想扭转这种局面，也要靠我们自己。

接下来的一章，我将展示遗忘在人类千年的历史中所扮演的角色，以及将记忆外化如何使记住对我们而言成为可能，甚至是历经几代人穿越长久的时间仍然能记住，但却从未改变遗忘作为常态的局面。

再下一章，我着眼于技术的发展——数字化、廉价的存储器、易于提取以及全球性覆盖，这些已经改变了记忆的经济学，并促进了遗忘的终止。

第4章将更加详细地描绘这种完整的数字化记忆对于个人与社会的潜在后果。而大家所提出的对策是第5章的关注点。由于这些对策还无法完全使我信服，我将在随后的一章中补充自己的解决方案。我的建议是，设置有效期限，通过记忆的有限性去面对信息，并且促使我们去理解和领会“信息也有寿命”。不过，我最为重要的目的是，想设法确保我们将能够记得在数字时代如何去遗忘。

---

[1] 寒蝉效应，指人民因害怕遭到国家惩罚而不敢公开发表言论，如同蝉在寒冷天气中噤声一般。——编者注

02 遗忘，人类的天性：遗忘的作用与人类记忆的演进

遗忘，是人类的天性。从古至今，人们不断尝试用本能、语言、绘画、文本、媒体、介质，来记住我们的知识。千年以来，遗忘始终比记忆更简单，成本也更低。而数字时代颠覆了这一切，我们惊愕地发现，如果真的记住一切，不仅令人发狂，而且让人孤独绝望……

大脑在不断地重构记忆——我们所记住的，有一部分是基于我们当前的偏好与需要。人类的记忆，是一个活跃演进的构造。

人类对记忆的构造与重塑，与其说是缺陷，还不如说是优点。利用归纳，依靠推测，强调当前时刻，考虑后来经历，这些都能够帮助我们敏捷有效地进行推论，进行抽象与概括，并及时采取行动，而非一直陷在相互矛盾的记忆中。

人们能够捕捉并记录的信息量，在最近的四分之一个世纪中有了大幅度的增加。但是在模拟时代，有效的记忆仍然既复杂又耗时，因此也很昂贵。在那时，记忆比遗忘要困难许多。

人类的本能记忆

我们所有人都经历过这种感觉：在聚会上碰到某个熟人，却无论如何也想不起来这个人的名字；在自动取款机前面，拼命回忆有一段时间没用过的银行卡密码；在停车场里四处徘徊，因为实在记不得具体把车停在了哪里。我对此深有感触，并总是埋怨自己在这些时候的记忆力太差。我们不喜欢总忘事，但是，遗忘，却非常符合人类的特征，它是我们的思维进行工作的一部分。不妨问问自己，还能详细地记得上周你跟伴侣之间的对话吗？两天前的晚餐你吃了什么？今天早晨电视上具体说了些什么？

人类的大脑极其复杂。它包含了一千亿个神经元，神经元是专门处理信息的细胞。每个神经元通过突触，与其他神经元存在上千个联结，以传递信息。这使得一个人的大脑中大约有一千万亿个这种神经联结。人们普遍认为，我们仅使用了大脑很小的一部分，而恰恰相反，健康人类大脑中神经元与突触的整个网络都是活跃的。但是，如果我们想要记住接收到的每个感觉刺激的话，大脑这种惊人的、巨大网络式的加工与存储能力将会被迅速淹没。

以视觉为例。平均而言，人类眼睛一秒钟能够捕捉几十张高分辨率的图像——相当于1500万像素的数码相机拍摄的照片那样清晰。在每天能够有意识的5万秒时间中，即便是非常大型的存储系统也将很快被装满。而且，再加上人类除视觉以外的其他感官，传入的信息总量将会更

加庞大。

为了应对外在刺激信息的汪洋大海，我们的大脑在将信息存入长时记忆（long-term memory）之前经过了多个层次的加工与过滤。比如，从几百万的脉冲中，我们的大脑构建出了一幅图像，可能是某个黑色的、会运动的东西，而瞬间之后我们“意识”到了有一只小黑狗向我们跑来。与某个越来越近黑色东西的一系列快照所包含的成千上万片信息相比，一只快速跑向我们的黑狗这个相对抽象的信息，则更为有用。更为重要的是，这个抽象的信息也更加简洁。在我们的神经元加工传入信息时，从简单的刺激到图像模式识别的过程，大量的信息都被有意地舍弃了。这是第一层无意识的生物学遗忘——所以我们很少能够意识到。

一旦某个外在的刺激通过第一个过滤层，或者此后我们形成了关于此刺激的意识，那么它通常会被存入所谓的“短时记忆”（short-term memory）中。在这种状态下，我们能够很容易地提取这种记忆，但是存储在短时记忆中的信息消散得非常快——大约几秒钟，大部分的信息就都消散了。

心理学家艾伦·巴德利（Alan Baddely）认为，短时记忆实际上包含三个独立的部分，它们都从属于一个特殊的控制单元。第一部分给视觉空间印象充当短时的心理画板。研究者已经表明，它能够保存几“块”（数位或者字母）信息，可能平均在4～7块之间。第二部分的功能是充当一个短时的音频回路，记录大约两秒的声音与言语。如果没有更新（比如，不断地重复），音频信息将很快消散。我们能够通过重复地念电话号码而短暂地记住它们一会儿，就是这个道理。第三部分是情景记忆（episodic memory）的一个短时缓冲区，情景记忆也就是对我们刚刚经历过的事情的记忆。

取舍之道

短时记忆是信息加工的瓶颈，几乎所有信息在到达长时记忆之前都必须通过这一瓶颈。一旦一片信息被转入长时记忆存储器中，它就被存储在我们所称的“人类记忆”中了。

实际上，记忆的过程分两步。第一步是成功地将信息转入长时记忆存储器中。第二步是从记忆中提取那个信息（见图2—1）。

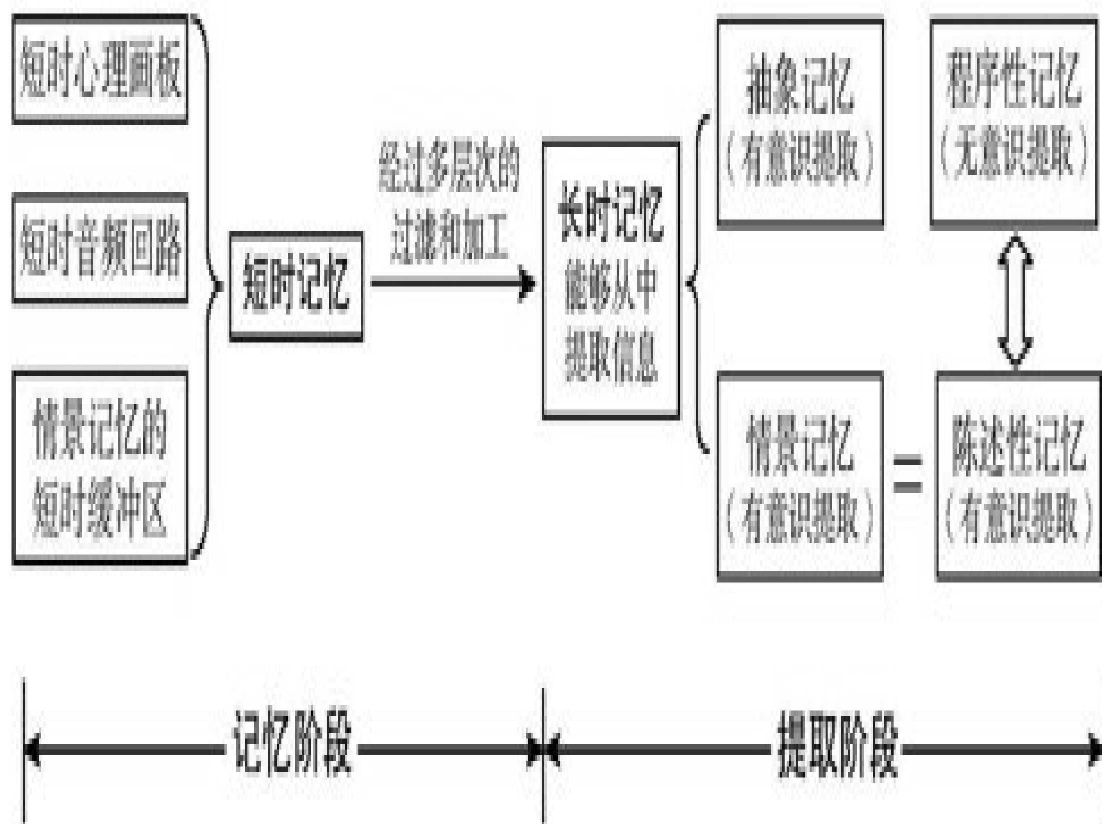


图2—1 记忆的类型与阶段

在人类历史早期，我们的祖先就已经发现，把某些工作重复多次将能够最终那些工作会存储记忆。比如，当早期人类的孩子第一次看到他们父母打猎时，随后他们就会自己尝试，一次次尝试之后，成功的狩猎过程就铭刻在他们的脑海中，当他们真正开始打猎时这些记忆就会被提取出来。收采蔬菜与水果、烹饪以及缝纫——这些都是我们祖先通过反复尝试学会的日常生活经验，而且他们后来能够很容易地提取这些经验。科学家称其为程序性记忆（procedural memory）。这种记忆并不是一种有意识的行为，而是人类从事特定日常活动时产生的副产品。这就是为什么它是内隐记忆（implicit memory）的一部分——内隐记忆是指我们在没有意识到这种记忆存在的情况下，能够获得并提取的一种记忆。在儿童时代反复多次尝试系鞋带之后，长大成人后我们自然而然地就记住了如何系鞋带。骑自行车、刷牙、或者翻筋斗，都是与这种经验类似地活动。

陈述性记忆（declarative memory）与程序性记忆完全不同。如果说从程序性记忆中提取信息是自动的，也就是说我们并不需要去牢记如何骑自行车，学会骑车后我们只是自然地骑上了自行车，那么陈述性记忆

则需要一种有意识的提取。比如，对于自己的初吻，一次重病或者一个特殊假期的记忆，都属于这类记忆。如果我们想要记住它，我们便不得不去想它，并且积极地在脑中去“搜索”，以提取我们过去的经历与事件。因为这些都是我们生活中已经经历过的特殊情景，所以它们被称为情景记忆。

情影记忆不同于抽象记忆——人类也具有存储并提取抽象知识的能力：记住数学中的勾股定理属于抽象记忆（abstract memory），而提取我们学会勾股定理的环境背景，比如高中，则是情景记忆。有时候，情景记忆与抽象记忆会混在一起。想象一下，观看莎士比亚戏剧《威尼斯商人》的表演时，你听到了夏洛克（剧中的犹太放高利贷者）说的台词。关于你所听到的，你记得的内容既是抽象记忆，也是情景记忆；抽象记忆是抽象的概念——演出的台词，而你经历过的具体情景则是情景记忆。

“遗忘”了存储在长时记忆中的信息究竟意味着什么，神经科学家与心理学家仍在争论。一些人认为，长时记忆中的信息是无法被擦除的，除非有生理上的损伤。他们提出，当我们遗忘的时候，我们失去的并不是信息本身，而是指向信息的链接。这就像互联网上的某个网页，没有其他网页与它相链接。没有链接指向它，该网页上的信息也就无法被发现，甚至通过巨大的搜索也无法被发现。这实际上就是，它被忘记了。不幸的是，对于提取记忆而言，这种如此关键的、指向我们长时记忆的链接，却并不是非常持久的。大多数链接会非常快地消失，这有效地剥夺了我们在长时记忆的海洋中访问信息的通路。

哈佛大学教授丹尼尔·沙克特（Daniel Schacter）对这种将人类大脑机械描述成一个巨大的、精密的、不尽完美的档案橱柜持怀疑态度。他认为，我们必须小心，不能陷入这种用数字计算机存储并提取信息的方式来类比人类大脑的隐喻，同时也不能落入现代信息处理理论将人类大脑概念化为确定性生物计算机的阴影之中。与这种机械的观念相反，沙克特教授提出了一种新的关于人类长时记忆的观点，即长时记忆并非不可改变地铭刻在石头上，而且我们也不仅仅只提取这部分记忆。沙克特提出，我们的大脑在不断地重构我们的记忆——我们所记住的，至少部分是基于我们当前的偏好与需要。对于沙克特而言，我们的记忆是一个活跃演进的构造。

实证研究似乎也支持他的观点。沙克特对关于人类记忆延展性的研究结果进行了分类，并半开玩笑地将它们归入他所定义的“记忆的七宗罪”当中。

第一宗是随着时间遗忘并非如最初所想的那样，遵从一个简单的衰

减函数。恰恰相反，遗忘取决于我们对于某个特定事件多久回忆一次，还取决于这个事件对我们的重要程度。第二宗同样也与此相关，我们很容易出错——认为我们记得某些事情，而实际上这些事并没有发生。通常，这种情况会与沙克特提到的“暗示感受性”相结合。暗示感受性，是指我们的记忆受外界作用所影响的程度。诱导性的提问是个非常典型的例子，看过法庭审判题材电影的人对此都会很熟悉。偏见会形成一宗相关的**“罪”**：它们会基于我们当前时刻所相信的，去歪曲我们关于过去的回忆。比如，如果在看到一只小黑狗之后，我们随后被问到狗的项圈是什么颜色，很多人将会自信地说出一种颜色，即便这只狗实际上并未戴项圈。总之，这些效应都反映出了一点，我们的记忆是不断构造与重构的。

如果沙克特是正确的话，是不是我们就会面临一种可怕的命运？即我们的过去不断被我们自己关于它的记忆所改变，以至于我们的过去变得短暂，能够自适应，而且还无法保证准确性。是不是人类的这种根本性缺陷，使得我们自从有人类以来便向往并努力去保存记忆吗？沙克特并不这样认为。他认为，人类对记忆的构造与重构，与其说是缺陷还不如说是优点。使用归纳，依靠推测，强调当前时刻，考虑后来经历，这些都能够帮助我们敏捷有效地进行推论，进行抽象与概括，并及时采取行动，而非一直陷在相互矛盾的记忆中。

不仅博尔赫斯，还有AJ，那位拥有惊人记忆能力的女性，可能都会同意这种观点。

美国加州大学欧文分校的伊丽莎白·帕克（Elizabeth Parker）与她的同事研究AJ多年。他们认为，AJ是一种罕见的拥有超常记忆力（记忆过剩）的案例，世界上其他被确诊的这类病例也只有少数几例。无需努力就拥有几近完美的长时记忆，听起来像是上苍赠予的礼物。不再会忘记钥匙与汽车放在什么位置，不再忘记生日与纪念日，不再有记不起老朋友名字的尴尬时刻，也不再会因为忘记已经看过某部电影而看第二遍。

然而，对于AJ而言，“记住一切不仅令人发狂，而且令她感到孤独”。与她有过深入交谈的约书亚·福尔（Joshua Foer）这样写道。她过人的记忆并没有使她更快乐；也没有让她在职业生涯中取得非常大的成功。她过着相对正常的生活，但却会花费非常多的时间陷入自己的过去当中，而非享受现在。

遗忘不仅对我们人类很重要，对很多其他生物也非常重要，甚至可以说对所有生命都很重要。实际上，记忆的困难可能是热力学第二定律的一个隐含结果。热力学第二定律是最基本的自然规律之一，它认

为，我们的宇宙（作为一个热力学系统）的随机性是必然会增大的，对此我们毫无办法。当然，我们能够人为地消除一些随机性，比如通过将气体分子放回到密闭容器中。但是这么做需要做出努力，用物理学家的术语来讲，这需要能量，如此便造成了比我们开始时更大的整体随机性。不是容器内部的随机性，而是容器外部的随机性。创建记忆是在我们的大脑内制造某种秩序，这也需要能量。另一方面，遗忘也可以是随机的，不需要高耗能的排序。因此，从根本上讲，物理学也告诉我们，记忆不像（随机的）遗忘，往往是有代价的。

取舍之道

几千年来，人类一直在尝试提高他们记忆的能力，增加他们能够存储并成功提取的信息的数量。与记忆相比，遗忘的力量更大，但是不难理解人们愿意孜孜不倦尝试的原因——因为记忆能够帮助我们更好地应对生活。

每一天我们都面临着多种选择，却无法真正知道这些选择将会导致什么结果。在决策过程中，我们使用我们过去的经验（情景记忆），和关于世界的一般事实、思想与概念（抽象记忆），以求我们所记住的过去能够弥补关于现在与未来的信息缺乏。出色的信息提取能力能够促进某个人的决策并因此增加他的生存机会。也可能是过度地敬畏这一点，我们的祖先长久以来一直欣赏具有惊人记忆能力的人：那些能够记住方向，或者太阳、月亮与恒星运动的人会受到欣赏。

然而，成千上万年来有两个障碍一直限制着人类记忆量的大小。首先，在存储并提取抽象信息方面人类一直存在困难。人们提取一个关于他们自己过去的事件都没有多大问题，如他们进行过的一次重要杀戮或者他们孩子的生日，但是记住某些像勾股定理一样抽象的信息仍然很难。其次，记忆被限制在每个个体当中。人们将记忆从一个人传递到另一个人的能力很有限。知识与经验，虽然它们对于他人可能具有巨大的用处，但是唯有克服极大的困难才能同时跨越空间的距离与时间的长河传递给他人。一位制造石刀的专家唯有通过培养另一个人才能将他的知识传递下去。程序性记忆的创建，个人感受以及经验铺就了缓慢而荆棘的传递之路，信息穿过这条通路时已经明显感到力不从心。

语言记忆

语言削弱了这个障碍。通过语言，人类能够彼此交流。从一个人传到另一个人的信息，接收者不必经历就可以将信息存储到记忆中：我可以告诉朋友往北走10分钟能看到一个小池塘，而不必抓住我的朋友亲自把他拉到池塘那里。语言还能帮助人类克服他们生命的有限性。在人类

拥有彼此交流的能力之前，生养孩子（即传递一个人的基因）是人类为子孙后代保留一些他们存在的唯一方式。语言增加了一种全新的策略，以战胜人必有一死的命运。

进化生物学家估计，在几十万年的人类进化历程中，语言是相对较近才出现的现象。只有当我们的早期祖先开始使用双腿直立行走，头部慢慢改变形状，大脑体积增加，声道改变形状，他们这才具备了适宜的生理基础以足够好地控制声音来说话。但是生物学上的进化并不意味着他们真正能使用它，同时由于缺乏直接的证据，因为我们并没有目击证人，或者当时的录音，有的只是间接的迹象。使得我们还不清楚我们人类的祖先何时开始说话。

关于大约生活在13万年前、善于使用石器工具的人类祖先尼安德特人是否发展出了一种语言用于彼此交流，以及他们的语言发展到了何种程度，考古学家与生物学家之间一直存在争论。目前的共识认为，在最乐观的情况下，它应该是一种超越了我们称为言语的类似歌唱的原始语言。但是在非洲，大约10万至5万年前，发生了一些事情。工具用途的复杂度与扩散度，膨胀得比以往任何时候都要快得多。很难相信，仅仅通过操作进行学习（“做中学”）就能够解释这种快速的创新与对新用途的采用，这是创建程序性记忆非常耗时的过程。而似乎更加合理的解释是，作为一种更为有效的沟通工具的语言出现了（从而传递了信息）。

取舍之道

我们的祖先迅速意识到，语言的作用非常大。它能让人们将他们的经验与知识传递给他人，从而能够很快在十里八乡进行传播。更为重要的是，语言使得人类能够通过将记忆从这一代人传到下一代人来保存记忆。这确保了新一代人能够建立起关于他们先祖的知识。

语言也改变了时间被人所感知的方式。在语言出现之前，上几代人他们的经验、想法与价值观念很少被人记住，只能通过形成程序性记忆的方式不甚明了地传给后人。有了语言之后，上几代人就成为了可确认的实体，他们能够被参考并且被记住。语言使人类能够了解历史，能够理解发生在他们之前的往事，并且最终将他们自己也贡献给历史。通过语言，在几代人之间传递的故事开始涌现。这些故事可能能够提供知识，或者仅仅只提供了娱乐，但是在这些早期的故事之外，还形成了关于人类的宏大叙事，正是这些人类的宏大叙事为生命与人类的存在提供了意义，并且这些叙事还传达了一种时间感与历史感。

所有宏伟的史诗都开始于口述传统：从苏美尔人的《吉尔伽美什》（*Gilgamesh*）[1]，到北欧人的古冰岛诗集《埃达》（*Edda*），再到古

希腊史诗《伊利亚特》与《奥德赛》，通常都采用押韵与韵律，以便于记忆。最重要的是，通过语言它们才得以形成。也是通过语言，人类才能够穿过时间的长河。

知识与故事这两种陈述性记忆的传递都是通过语言促进的。当人类使用语言进行交流时，彼此的经历就能够被分享。不需要人们共同参与某个具体的过程，情景记忆就能够被创建出来。告诉一位朋友我如何抓住一只鸟，比打猎时带上某个人让她观看或者为她抓一只鸟要快得多。这是个巨大的好处。另外，语言的影响在思想大意、事实与洞察力的传递方面更为明显。如果我想告诉别人我抓住一只鸟的经历，至少我还可以向别人演示我是如何做到的。但是，古希腊哲学家毕达哥拉斯怎么能够在没有语言的情况下传达他深刻的见解呢？

由于语言使沟通更便利，同时促进了抽象记忆的创造与传播，它改变了我们祖先的生活。语言使得他们能够进行抽象的思维，并且发展出对生命与自然的普遍观念。不同于他们之前世世代代的先人，掌握了语言的祖先有了发现这些思想、理解自然法则的动机，因为他们终其一生构建的思想见解、抽象记忆能够通过传播而保留给下一代。

结果是惊人的。语言释放了人类个体记忆的潜力。正如J.R.安德森（J.R.Anderson）与威廉·麦克尼尔（William McNeill）所认为的那样，这使得比以往大得多的人类群体间配合成为可能，语言促进了更大社区的形成并促使更大群体中的人们一起劳作。语言培养了对知识的运用，并释放了创新与效率的力量。总之，语言对人类发展的影响，与火的发现同样重要。

尽管说交谈很便宜，但是通过语言使人类记忆代代相传却并非不需要代价。实际上，代价还相当高昂。它需要时间与精力去传播，去记忆。年轻人与年长者需要聚在一起，年长者要去讲述，同时年轻人要去倾听。讲述者必须准确地传达事件的本质，而接受者必须尽可能多地去集中保存其本质。

取舍之道

极有可能，为了增加回忆的准确性，这些故事不得不多次地被讲述，以求完全被理解，这使得语言记忆既耗时又昂贵。因此，一个事件被记住所需要的努力就成为了一种有效的过滤机制。最微不足道的见解或许就不会通过这种方式去传递；这一方法是留给我们祖先认为的最重要的东西：那就是能改善人类生存的知识，以及思考了关于人类存在的根本问题，并帮助在人群之中与几代人之间形成了共同纽带的宏伟史诗。

今天，我们拥有令人印象深刻的技术帮助我们彼此之间进行沟通，

从移动电话到视频聊天。然而，这些科技设备并没有改变其根本：人类间的沟通仍然是耗时的，所以，人类要记住其内容也是如此。

直接的口头交流，是我们人类用来提高自身记忆能力的一种机制，但是它有其内在的局限性。它只有当我面对一个能听见能理解的同类时才有效，如此才能够将我回忆起来并告诉她的见解与经历添加到她的记忆中。如果我面对的是不能交流的对象，或者我周围的那些人并没有兴趣听我讲，将会怎样呢？尽管拥有沟通的工具，但是如果我无法分享我的记忆它们依然会消失。分享是保持记忆鲜活的关键，而语言是分享的一种主要机制。

不幸的是，仅仅分享还不够。即便有人在听我讲我的记忆，我可能无法准确地回忆起来。正如我之前解释过的那样，人类的记忆力相当不错，但它远非完美。记忆一直是一种长期的建设性的努力。通过回忆，经历与思想再度出现在我们的脑海中，但在我们分析并存储我们听到的内容时，会受到我们自己个人发展与我们生活环境的影响。这就像孩子们玩的“土电话”一样，传话的效果只会越来越糟。当孩子们细声耳语，一个人一个人地往下传时，他们在传给下一个人时已经对听到的话“添油加醋”了。这导致的结果就是，队伍中最后一个小孩说出的话，往往与最先传出的原话风马牛不相及。

在许多方面，几代人之间口口相传的口头记忆就像是玩“土电话”，不过只有一点不同，那就是我们并不知道原始的故事是什么样的。在两个方面，它实际上比“土电话”还要糟糕：首先，我们通过几代人之间的讲述与记忆所传递的东西并不是一句简单的“你好”，而是一组复杂的信息。虽然在某种程度上，这可能会使准确记忆的可能性更大。一个详尽的、扣人心弦的故事提供了丰富的背景，而人类发现这背景比脱离了具体意义的一个词或短语更容易记住。因此，人类记忆与口头交谈的结合可能能够保存一个故事的核心要点，以及生与死、爱与战争的惊险史诗。但是另一方面，通过这样一个过程不可能长时间地保存准确的细节。这种细节可能对于一部史诗传奇来说是次要的（比如，德国民间史诗《尼伯龙根之歌》中的英雄人物齐格弗里德杀死的龙是绿色的还是褐色的），但是，当准确性非常重要时细节就起到了重大的作用（比如，如何种植农作物，对天气的解释，依靠星星航行，或者使用某种药草治疗某种疾病）。这使得我们的祖先非常珍视那些拥有准确记忆能力的人。

“拥有良好的记忆被视为卓越的标志；它在找工作与成为社会精英方面有极大的帮助。”美国纽约大学英语系教授玛丽·卡拉瑟斯（Mary Carruthers）如是说。柏拉图与亚里士多德，以及其他希腊哲学家充分

地阐述了准确回忆的价值与本质，并描述了多种改善一个人记忆力有用的工具。人类社会对于出众记忆能力的崇拜一直持续到中世纪，并坚持到了现代。

然而，人类记忆的基本特性一直没有变化，那就是：存储与回忆是一种需要不断建设的过程。如果我们想要使用语言将我们的记忆传递给他人，那么事情就变得更加困难了。这时候存储与回忆的建设性过程就发生在两个或多个人之间，而不是某一个人身上。词语的意义以及它们的内涵在不同人的理解中是不同的，而且那些只听过故事而没有经历过的人，不得不依据他们对听到的词语的理解，完全按照字面意义去构建记忆。尽管拥有这种结构性的不精确，孩子们用“土电话”传递故事时所领会的东西，还能够给他们带来一种类似在一群人中建立起一份共享的公共记忆的感受。但是，通过人类记忆的每一次提取，倾听与存储的内容已经被改变了，尽管变化是如此轻微。

外部记忆：绘画与文本

如果人类自身的记忆是一个薄弱环节，那么可能记忆就需要从人类的大脑转移到一些外部存储器与检索设备上。通过绘画或者书写，我们捕捉到一个事件、一种情感，或者一种思想。看看我们自己画的画或者阅读我们自己写的文字，能够帮助我们记忆，使得我们更准确地回忆过去成为可能。这样，外部记忆成了我们人类自身记忆的一种延伸。

外部记忆也能够被用于促进共享的公共记忆的构建。比如，当我写了一封信同时其他人又读了这封信的时候，这种记忆就产生了：我书信中的内容就成了阅读者与我之间共享的记忆。当人们通过彼此交谈来分享经历时，他们同时也在构建一种共享的记忆。然而，在这种情况下，讲述与倾听是同时发生的，而未必需要在同一地点——如果他们使用电信设备的话。相形之下，有了外部记忆之后，讲述与倾听就成了两个在时间上可以分离的独立事件。当我写一封信时讲述发生了，当收信人读到信时倾听也发生了。讲述与倾听共同构建或重建了共享的记忆。

绘画

绘画可能是人类构建外部记忆最古老的形式。它创造出一幅关于某个场景或事件的图像，不管是真实的还是想象的，这样便能够使人记忆。人类已知的最古老的绘画可以追溯到三万多年前。它们在世界各地的洞穴里可以看到，主题常常是动物与猎人。随着时间的发展，绘画被用于记录历史事件，从而保存了人类的经验。

最初，绘画与超自然现象联系在一起。在洞穴里画过动物的人们，

希望通过画它们能够使它们出现。法老和他们的家族希望绘画能在他们死后将他们带到神灵的面前。宗教的动机在希腊与罗马的绘画中继续扮演了重要的角色，当然在基督教兴起到进入启蒙运动期间也起了重要的作用。

同时，绘画也记录了社会中最重要的事件：战争的胜利或者失败，社会的发展与财富，诸如洪水、干旱、火灾之类的自然灾害，或者广泛流行的疾病。上层的贵族经常被突出地描画，并交织出现在绘画叙事中。

随着宗教对绘画支配性的减弱以及人们接触艺术与艺术家的机会增多，绘画的功能已经超越了保存社会的重要时刻，转而去捕捉一个更为广泛的群体的观感。绘画变成了一种创造个体记忆的途径。艺术化的自我表达，强烈的宗教与精神潜流都是绘画的重要动机，此外，还包括捕捉并保存某个永恒瞬间的能力。华美而精细的埃及坟墓绘画描绘了已故之人的生活场景，赫库兰尼姆[2]亮丽的镶嵌画与庞贝古城[3]的壁画，以及吴哥窟庙宇中令人震惊的浮雕描绘了整个人类的历史，这些都是艺术保存人类记忆的最知名的证据。

这种艺术化的记忆当然也要付出代价。它不仅耗时而且昂贵。资深的艺术家不得不苦干很长的一段时间，才能创造这些令人印象深刻的艺术作品。在古巴比伦，古埃及，古印度，古中国，古希腊与古罗马，绘画只有相对较少的精英才能接触到，画作记录的也是他们的一些个人经历。通过文艺复兴时期财富的传播，尤其在城市中，欧洲成功的商人家庭能够支付得起一位画家去捕捉他们生活的时刻——但是由于画家需要耗费好几个月，甚至好几年才能完成委托给他的画，所以绘画仍然非常昂贵。又过了两个世纪，着色的家族肖像画才变成了中产阶级一种创造记忆的常见方式。即便在那时，绘画也绝不会很廉价，人们都是小心翼翼十分爱惜地使用。今天，绘画（比如雕塑）仍然是一种昂贵又耗时的外化人类记忆的方法。

通过绘画保存记忆还有另一个更为严重的缺点——它善于捕捉某个永恒的瞬间，但是却不善于捕捉时间本身，因此它在记住一个故事，一个人，或社会生活中的一个插曲方面，作用很有限。一个人能够很容易地想象出一幅关于战争中决定性时刻的画作，但是一位画家要描绘整个战争如何展开就没那么简单。当然，一系列的绘画能够反映故事中的重要时刻，但是它却是一个更加耗时的方法，而且仍然需要将故事的大部分留给观看者去想象。

取舍之道

时间并不是绘画难以表现的唯一维度；抽象的观念或者思想，比如勾股定理或者牛顿的万有引力定律，都很难用绘画去表现。

当然，一个人能够想象用一个画有各种三角形的漂亮插图去传达勾股定理的本质，并想象出用各种尺寸的下降物体去传递牛顿原理的核心，但是这种绘画将会产生很多种不同的解释，可能会造成令人无法忍受的歧义性。

文本

由于人类记忆的不精确和绘画固有的缺点，我们的祖先便去寻求不易修改又没有多重解释的外化记忆方法。尤其是那些专注于生产、贸易与管理的组织，想要拥有一种能够简便精确地存储与提取信息的方法。这就有了文本的诞生。让人非常诧异的是，古代的官僚主义与会计人员正是促成因素。在公元前四千年，居住在肥沃的美索不达米亚的人们已经超越小的村庄形态，形成了更大的群体。因此，这种早期的社会就需要超越了通常确保作物生长，物资交换与社会管理的组织。正如考古学家丹尼斯·施曼特·巴塞瑞特（Denise Schmandt Besserat）告诉我们的那样，为了记录账目与存货，便有人用铁笔在粘土上标记小的凹痕进行记录。这些早期的象形文字，最终逐渐形成了书写，使公元前三千年早期苏美尔人的楔形文字成为第一种书面语言。书写大约在同一时期也在埃及出现了——埃及象形文字，在几个世纪之后在中国也出现了文字。

这对记忆是一个突破。一旦书写被大家所知并确立下来，人类的经验与知识就能够被存储在人类的头脑之外，并能够随意准确地进行提取。它使记录复杂的灌溉与种植业成为可能，并进一步促进了广泛的贸易。书写促进了国家的诞生，因为它使征税，以及对军事和民事信息权力的管理变得简单。对自然的观察与工程学的洞察力现在能够被保存、被分享，并能够准确地提取，而且技术发展的传播使得复杂的结构与了不起的建筑成为可能，比如埃及的金字塔，雅典卫城，或者中国的长城。但是书写的出现并没有改变一个事实，那就是记忆仍然是需要不断构建，耗时又昂贵的。生产早期的苏美尔象形文字是一件非常费力的事，因此只能非常慎重地使用它们。苏美尔人耗费了几百年时间才将精心制作的象形图，归纳为更为程式化（因此耗时也更少）、代表不同词语的经典苏美尔楔形文字。大约2000个不同符号需要一个非常有经验并受过高度训练的抄写员。这些过程很难绕过，这使得记忆的代价依然高昂。将大量的符号整理并简化成大约300个常用的文字，差不多花了两千多年，从而使得书写变得更能够负担得起。

之后，楔形文字被一种更灵活更有效的书写系统所替代：那就是字母。通过采用代表音素的符号，比如元音或者辅音，去替代完整的单

词，一个抄写员不再需要记住几千个或者几百个不同的符号，而只需要记住26个字母。它大大降低了文字的复杂性与抄写员所需要的训练，最终增加了它们的供应并降低了书写的代价，降低的代价可能高达一个数量级。

即便使用的是繁琐的楔形文字，抄写员（与他们的老师）书写的内容也远远超出了账目与官僚记录这个狭窄的领域，而他们原本正是被雇来做这个工作的。就这样，文学作品出现了，比如史诗，寓言与散文；其中一些是虚构的，但是许多作品也试图记录真实的事件，经历与情感。这些情景记忆的样本现在通过外化而被人们所记住，并保存在泥板文书和纸莎草纸上。

书写对于保存抽象的观念与思想更为有用。数学与科学知识首先被人类所记录，公元2400年前政治与历史领域的知识紧随其后，同时期被记录的还有第一部法律。

书写也成为了一种信息权力新的来源：这种信息权力就是保存知识。领袖们开始出于他们自己的目的，对书写作品的收集进行制度化。亚述帝国的末代国王巴尼帕（Ashurbanipal）可能是第一位正式被训练成抄写员，并且能够说多种语言的领袖。在他追求创建记忆与知识贮藏室的过程中，他将他征服了的城市中的粘土块与纸莎草纸搜刮一空。整洁地存储着的粘土块上成千上万的文本，可以想象，亚述帝国的图书馆一定是一个令人惊叹的机构。

几百年之后，另一个创建了广泛外化记忆收藏的工程获得了世界性的知名度，这个工程是由统治亚历山大港（今埃及）的托勒密家族实施的。在其鼎盛时期，它是世界上最大的图书馆，估计藏有五十万册著作。不像亚述帝国，亚历山大港的图书馆收藏的主要是纸莎草纸卷轴，这种卷轴与粘土块相比重量轻，而且能够被卷起来，因此可以节省空间。托勒密家族对此工程非常专注，并且在他们追求增加收藏量的过程中非常残酷无情。他们诱使精英知识分子从遥远的地方来到亚历山大港，让他们进行研究、书写，并将自己的一生最终贡献给他们的图书馆。到访亚历山大港的访客都被要求放弃他们所拥有的所有书面材料，并交到图书馆。国外所持有的重要文本，都以非常高昂的代价借过来并进行复制。

曾经，托勒密三世用一个巨额的抵押品去借雅典人的一些最重要的文件。他小心翼翼制作了非常漂亮的副本，并将这些副本“送还”给雅典人，将真品留了下来。他说雅典人可以自己留下那些抵押品。

托勒密三世这么冷酷是有理由的。他认为，一个文本的每一个副本总是会产生错误，并且这个错误最终会降低它的价值。这就是为什么他

更加喜欢获得尽可能最古老版本的原因，他希望那个最古老的版本将能尽可能地接近于原始版本（或者可能甚至就是原始版本）。在适当的时候，这导致了伪造“真品”的繁荣。托勒密三世的政策使得几千年前就开始出现的一个过程变得非常明晰，那就是：对于一个民族全面、几近完美记忆进行支配的渴望。

然而，与文本从根本上改变了我们人类保存信息、增强回忆的能力一样，它也证明了通常意义上外部记忆的固有局限，尤其是书面记忆的局限。这是因为，对于我们人类而言，随着时间的变化基于我们对世界不断进化的理解，词语的意义发生了微妙改变，即便当初书写与后来阅读的是同一个人。

如果你曾经尝试阅读过你自己多年前写的一本旧日记，你可能会感受到这种熟悉与陌生相混合的奇怪感觉，你感觉你记得一些，或者可能记住了大部分，但却从未记住所有文字的原始意义。由于词语不是信息的精确载体，每个词语与句子都需要人类的解释，需要我们以我们回忆它时所理解的语义去解释它。文学理论家们正好就在争论，我们阅读一个文本时，它的内容（与它被写下时相比）改变了多少。但是他们的确也同意，阅读（以及对外部记忆其他形式的回忆），就像记忆一样，将需要时常在心中作出构建或重建的努力。

托勒密三世可能会认为，通过收集原始的文本，他就可以保有对原始语义的控制，但是实际上他并不能。只有通过与他实际上无法控制的两种行为——阅读与解释，只有这两种行为相结合，才能完全达到他在收集文本信息时想要证明的价值。

托勒密三世与我的一个法学院的朋友一样，都掉进了一个相似的陷阱。我这个朋友给他自己强加了一个有趣的系统，以提高他的外部记忆。他明确将他个人图书馆的藏书限定在200本以内。一旦他阅读了一本新书，他就要判断这本书是否属于他已经读过的最好的200本书的行列。如果属于，他将会把这本书加入自己的收藏，同时把其中不那么好的一本剔除掉。久而久之，他认为这个不断筛选并且有选择地进行遗忘的方法，将能够不断提高他图书馆藏书的质量。他要保留在外部记忆中的将只会是真正重要与有价值的思想。

然而，他未能了解的是，由于他自己的观点是随着时间变化的，所以他评价并决定哪本书该留哪本书该丢的基础也在不断变化。经过几十年的勤勉筛选之后，他的图书馆收藏的将不会是他曾读过的最好的书，而是代表他生命中不同观点、价值与阶段的书。就像托勒密三世一样，他所忽视的是，根本不存在一个随着时间变化不会改变的客观评价，进而我们人类也就无法用它来判断我们阅读的文字，或者我们回忆起的外

部记忆的意义与价值。

在古希腊与古罗马，抄写员这个职业坚持要求专业人才。这个职业要求昂贵而旷日持久的专门训练，并且对于绝大多数人而言无法企及。但是在富人中间，书写却在传播。比如，在古罗马的庞贝古城中发现的镶嵌画与绘画中，描绘了一对有学问的夫妇，丈夫拿着一个卷轴，妻子则拿着一支书写用的尖笔。渐渐地，卷轴被抄本代替了，在两块木板中间有一组板片绑在一起，已经跟现代的书籍很像了。卷轴的本质是倾向于维持文本按顺序展开以及叙事的上下文背景，而有了抄本之后，读者至少在原则上接触文本的某一部分要比以往容易得多。

图书馆在继续繁荣发展。图拉真皇帝（Emperor Trajan）的罗马广场不仅仅是一座巨大的图书馆，而且还为其读者提供了一个阅读与工作的场所。图书馆还突然出现在很多小型的城市，以满足居民们想要接触关于人类的全面记忆的渴望。最重要的是，图书馆对于大众而言变得可以企及。亚述帝国末代国王巴尼帕那所最令人印象深刻的图书馆，主要是供他个人以及他最信赖的仆人使用。相比之下，罗马人将图书馆建在公共浴室旁边。识字的市民现在既可以洗澡也可以阅读——这同时反映了图书馆所有者对放松身体、缓解焦虑、传播知识的需求。

然而，书籍仍然非常昂贵。它们都是万中选一的；每册书都是经过训练的专业人员辛辛苦苦用手工制作的。最勤奋的抄写员在他们一生之中也仅能制作几十本书。

在公元11世纪的英国修道院里有一组抄写员，他们在22年里持续不断地工作也仅仅只复制了66本书。而且几乎同时，本笃会（Benedictine）的抄写员欧特罗（Otloh）据说在他40年的工作生涯中总共复制了28本书——产量还不到一年一本。当然，这些数字也不能仅仅照字面去理解。一个人不能根据少量的数据点去估计一位抄写员的平均生产力，尤其那时候的原始资料是众所周知地不可靠，而且许多手工制作的书籍包含丰富的插图与昂贵的封面，这些都需要耗费更多额外的时间。但是即便我们将这些因素都考虑在内，它们的影响仍然微不足道，复制书籍的代价依然十分高昂。据施拉姆（Schramm）估计，那时候为一本书制作一个副本，图书馆的花费要比今天的一万美元还多。即便最简短的书，一位抄写员全职工作也可能需要三个星期，售价大约为700美元。这种形式的记忆是一种非常昂贵的手段，使用时需要非常小心。由于遗忘才是常态，所以人们在文本出现以后继续遗忘。这种状况在接下来的一千年里仍未改变。

在公元第一个千年的中期到晚期，对于大多数人而言记忆变得更加昂贵。因为天主教会开始侵占大部分有文本的环境。在修道院中，阅读

与书写传授给了修道士，他们随后被教堂雇佣担任抄写员。这些修道院还具有图书馆的功能。总而言之，教会允许教堂去决定哪些书将被复制并被阅读，以及由谁去复制与阅读。这与自由的（虽然也有社会分层）罗马时期形成了鲜明对比。在这一千年中，记忆被严格控制并被有组织地集中起来，虽然“无声”阅读（而不是修道士大声的朗读）的传播，使得阅读对于那些被允许的人更有效率。书籍作为外部记忆的作用甚至连富人都无法企及：他们无法进入制作书籍的基础设施，也无法接触到阅读（或书写）所需要的训练。世俗机构也在挣扎，而教会对阅读与书写的掌控也仅仅只是缓慢地放松。在公元15世纪开始之前，剑桥大学的图书馆总共只有122本藏书，而经过半个世纪的努力这一数量增加到了330本。

到了大约公元1450年的时候，一个相对较简单的发明与社会学因素相结合促成了重要的改变：那就是德国人古登堡发明的活版印刷机。它使书籍的大量生产成为可能，大大缩短了复制一册书所需的时间。

在1483年，正如历史学家伊丽莎白·爱森斯坦（Elizabeth Eisenstein）所写，的黎波里出版社出版了《柏拉图对话集》（*Plato's Dialogues*）。用以往一位抄写员制作一本书的时间，出版社大量炮制，共生产了1025册书。而当的黎波里出版社对每一页的索价提高到抄写员索价的三倍时，它仍然令人惊愕地将劳工的工资压缩到前印刷机时代的劳工工资的三百分之一，而前印刷机时代才仅仅过去几十年而已。

生产革命像野火一样在欧洲传播。到1480年，印刷机在欧洲的110多个城镇中运作，这差不多是意大利一半城镇的数量。印刷术如此迅速激增，是因为几乎没有任何人能够阻止它。专利那时还不存在，而且古登堡自己又破产了，这就使得欧洲的其他人仿造了他的发明。更为重要的是，天主教会并没有控制古登堡的发明造成的影响，最开始教会甚至拥抱了这项发明。教会没能阻止住印刷术的传播，而正是印刷术，直接破坏了天主教会对书籍的控制。正如爱森斯坦提醒我们的那样，在1453年到1503年这50年间，大约800万本书被印刷出版，可能比1250年前君士坦丁堡建城以来“欧洲所有的抄写员制作的书籍还要多”——产出的书籍增长了令人瞠目结舌的25倍！印刷厂大批量生产娱乐作品，以及希腊与罗马经典书籍，还有相对较少一些的宗教典籍。到此时，对记忆的控制这才脱离了教会的手掌。

仅仅几十年之后，教会就彻底失去了对书籍的掌控。在1517年，相对而言不为人所知的马丁·路德（Martin Luther）用印刷机发表了他对天主教会的批评，开始了一场塑造整个欧洲大陆的广泛运动，促进了日后的“三十年战争”，并导致了新教的建立。到1521年，路德已经将拉丁文

的《圣经》翻译成了德国本土语言。德语版《圣经》次年出版，并且很快便成为一本畅销书。到1574年，一位出版商已经将路德翻译的《圣经》出版了十多万册。在那珍贵的几十年中，印刷术革命的进展几乎到了无法控制的程度。政府刚刚开始从教会的压制下将他们自己解放出来，他们也看到了削弱教会权力的潜在好处。最为重要的是，人们想知道，怎样才能不依赖昂贵而且腐败的中介机构，进入天堂。人们希望能够拥有直接走进上帝思想的本地语言版本《圣经》，那可以成为人们最珍贵的记忆。这一期望驱使人们去学习如何阅读。

《圣经》也被人们在家中与小型社区中大声朗读，这扩大了《圣经》的影响力。到16世纪末，欧洲已被各种各样的出版物与小册子淹没，其中有许诺能够让读者接触到上帝的，有促进宗教与政治宣传的，有记录科学发现的，有复兴古希腊与古罗马经典的，还有提供消遣与娱乐的。一个重要的转变已经发生：外部记忆已经变成大批量生产了。

然而，记忆从根本上讲依然昂贵。当然，印刷机的发明让劳动力的成本已经大幅下降了，而且随着天主教会失去其对书面语的掌控，对书籍的需求有了本质性的增加。但是，尽管出版从精心制作每一册书籍的抄写员，转移到能一下印刷上百或上千册的出版商手上，单本书的价格却没有下降很多。因为书籍是被印在纸上的，而这种关键资源的成本一直很高——这种状况持续了几个世纪。在1483年，当黎波里出版社出版《柏拉图对话集》的时候，他们为印刷这一工序索价90弗罗林（一种货币）。纸张的费用在120到160弗罗林之间，这构成了总生产价格的三分之二。费夫尔（Febvre）与马丁（Martin），在长期观察纸质印刷的价格后，他们得出结论称，几百年间它的价格几乎没有变化。真正进入18世纪中叶的时候，就著名的《百科全书》的出版而言，总共105里弗（古时的法国货币）中的68里弗（65%）都花在了纸张上。即便有了古登堡的印刷机与宗教改革运动带来的变化，对于大多数人而言，使用书籍仍然是一种难以理解的增强他们记忆的途径。

这一事实在欧洲以外也并无不同，虽然原因有所不同。在伊斯兰教中，印刷品（取代抄写员誊写的书籍）被视为亵渎神明之举，是在模仿神的荣光，因此被禁止。在中国，印刷机被人所知已经有几个世纪了，虽然他们使用的是木版印刷，而不是古登堡灵活高效的活版印刷机。[4]中国人一直在使用表意语言，比采用字母的表音文字更难以使用；而且高度分层的中国社会缺乏给16世纪的欧洲带来改变的那种张力。朝鲜人也开发出了活版印刷机。然而，与古登堡的发明不同，它一直是朝廷与精英的工具，他们将出版物的主要内容限制在儒家著作方面。

19世纪早期的工业化纸浆厂最终降低了纸张的成本，导致纸张生产飞速增长。在英国，1861年至1900年间，纸张的产量增长了7倍。就像蒸汽压熨机以及工业革命的其他先进技术给经济生产带来的效率一样，这使得出版商能够印刷比以往任何时候都要多的书籍——据估计，有18世纪书籍的4倍之多。随着书籍数量的增加，书面的外部记忆也在增加。

不断增多的书籍确实有人读了——不过在19世纪的头十年，却只有一小部分书籍被人买去阅读。有两点原因造成了最开始的读者人数如此之低。

第一个原因是读写能力。直到1840年，英格兰与威尔士的识字率才达到68%。这一情形与欧洲其他国家类似。第二个原因是书籍的价格——至少在有些国家是如此。大规模生产与更廉价的纸张本应该降低了书籍高昂的价格。但是正如西蒙·艾略特（Simon Eliot）详细阐述的那样，至少在英国，在整个19世纪，“书籍的价格竟然都维持恒定没有下降”。书籍依然很昂贵。对于富裕的伦敦律师而言，买一本书是一笔相对微不足道的费用。但是对于一位工人而言，购买一本书意味着他几个月的可支配收入。如果再加上晚上读书照明所必需的不菲代价（一般人在白天的大部分时间里都需要工作），阅读书籍不仅对于工人仍然高不可攀，对于大部分办公室职员与其他职业人员而言也是如此。

自16世纪初，期刊就已经出现了。到了17世纪，期刊已经随着富裕的专业人员与商人阶层的成长而流行起来。期刊还有一个优点，那就是它很容易在读者之间进行共享，比如在欧洲的很多咖啡屋里，这创造了公共领域的开端。但是读写水平太低，同时报纸价格仍然太高，因此期刊仍旧无法在欧洲真正成为一种大众现象。随着更便宜的报纸的出现，这一局面有了改观，其中包括所谓的便士报，它的价格大约只有传统报纸的六分之一。读写能力较低的中产阶级与工人阶级终于能够承担得起定期印制的报纸信息的价格，而且他们也开始这么做了。

在美国，这一情形发生改变的时机要早许多。普林斯顿大学的社会学家保罗·斯塔尔（Paul Starr）在他关于媒体兴起方面的权威之作中写到，到19世纪早期，报纸“在美国比其他任何地方都更常见”，而且“社会的识字率异常地高”。美国这个年轻国家的领导者，将促进一个既强劲又广阔的交际空间视作核心目标，以支撑美国的经济社会发展，而且美国国民也愿意拿出时间与资源来了解每天发生的事情。

除了识字率的提高，另一个明显的变化是，引入了廉价的通俗小说，这满足了低端市场的需求——至少满足了那些不太富裕的阅读群体的需要。但是，低端并不必然意味着廉价；这些通俗小说的价格美国民

众完全可以承受（因为相比其他许多消费品，书籍存在一个有利的税收制度）；但是在英国，即使廉价的通俗小说也相对比较昂贵，需要花费一名工人一个星期的可支配收入。仅在20世纪初，当社会改革开始改变工薪阶层的境遇时，他们的生活水平才得以提高，工作时间才开始减少，书籍和报纸的价格才大幅下降，最终整个英国的阅读人群才迅速扩大。

共享记忆

这一变化导致了外部记忆中的一种新现象——书籍能使一个人瞥见这个世界的一角。类似地，报纸以每日快照的形式向读者描述这个世界；长此以往，书籍和报纸向人们传递了一种过去每一天里世事如何变幻的感觉。在19世纪，不算便宜但还负担得起的书籍和报纸使人们广泛地共享记忆成为可能，即便这种对大事件的共同记忆的确存在偏见，即便这样传播的事件与以往发生的任何事件都不尽相同。当然，宗教也提供了一种共享的方向感，但是这种共享主要体现在信仰上。书籍则一直在为读者（包括每本书的作者自己）构建共享的共同记忆中起着某种作用。不过一般来说，这种共享的广泛性还不足以达到覆盖所有人群的程度。

如今，价格低廉的书籍和报纸同时做到了这两点——它们既转播了关于真实事件的事实，同时又进行了广泛的传播。从这种对大量印刷出版物的大规模阅读中显现出来的，则是广泛共享的社会记忆。通过阅读的行为，这种共享的记忆不断增加，为读者提供了一种地理上以及时间上的位置感。

首先，在地理上，读者能够让自己置身于与世界以及与世界大事件发生地的关联中，不仅超越了地理位置的限制，而且正如本尼迪克特·安德森雄辩所说，还为自己提供了一次属于一个强大的、即便是存在于想象中的共同体的机会。在这个共同体中，人们感到彼此相连不是因为他们在地理上接近，而是因为他们已经在社会层面建构了，或者说是想象出了，一个可以归属的共同体。对于安德森而言，这将大量印刷的出版物（他称之为“印刷资本主义”）的兴起与帝国和民族国家的崛起联系在一起，而无论帝国还是民族国家都属于想象的共同体。

同样重要的是，完整共享的记忆也为读者提供了一种时间上的关联感。这对记住这些共享的记忆具有深远的意义。随着今天的新闻报纸在隔夜之间变成旧闻，今天的新闻事件就成为了过去。对于报纸读者来说，这些过去并不是一个隐藏的秘密，在最初读过之后就再也难以记起，而是他们自己在过去的每年每月每天里读过的东西，以及他们如何

记住这些东西。

读者通过每天读报的行为，直觉地抓住了这些消逝的过去，最终一次一个片段地构建起了共享记忆的印记。这是外部记忆与读者自己建立起来的主观联系和内涵之间的结合，也就是出版物上印刷的内容与读者个人理解并参考的内容之间的结合。

通过阅读报纸和书籍，人们能够记忆。比如，将珍珠港事件作为一个共享的历史事件记下来，即使事件发生时很少有人在珍珠港。类似地，人们可能感觉他们“知道”肯尼迪总统遇刺事件、美国东海岸大断电事件、2000年扣人心弦的总统大选，“9·11”恐怖袭击事件，并享有对这些事件的一份记忆。大众传媒培养了这种共同的共享记忆，而这种共享记忆既超出了人们共同目击的事件或者一个目击者告诉朋友的事件本身，也超越了地理狭隘的限制。

因此，对于19世纪和20世纪的报纸阅读者而言，即使很遥远的事件不仅能够在地理上，而且能够在时间上产生超越——同时一般都能记住。对于这种共同记忆的形成，必要的条件不仅仅包括承受得起的大众传媒的大量供应，以及人们同样巨大的需求。这还不够，由于人们阅读的书籍千差万别，这导致仅在作者与特定书籍为数不多的读者之间形成了共享记忆。只有当大部分人都读了同一本书和同一份报纸的时候，人们才会从相同的信息源中构建出共享的记忆。这不仅仅适用于报纸，也同样适用于广播和电视。当国民仍然只能收听数量有限的联网广播，或者只能收看有限的电视台时，广播和电视的晚间新闻节目中播报的新闻就成为了我们社会共享的共同记忆的基础。著名新闻节目主播沃尔特·克朗凯特可以在每次节目结束时说“事实就是这样”，因为大众将以这样的方式记住它们。

至于这种广泛共享的共同记忆是好是坏，各方看法不一。一些人认为，随着人们都看相同的报纸，收听相同（或非常相似）的广播，通过对世界的共同理解，人们能够增进彼此之间的交流。对于另外一些人而言，在多样性如此有限的信息基础上，人们形成的对大事件的记忆将会非常狭隘。

不过双方都同意一点，那就是，对高度受限的信息供应类型如此集中的需求会限制信息的传播量，而正是信息形成了共享记忆的基础。如果大多数人都只看一种报纸，或者只看一个电视频道，那么得以传播并最终进入共享记忆的事件只会更少，而不是更多。最终结果就是，在出版物与广播市场集中化的传统世界，获取外部记忆依然很昂贵——对于接收者而言，这种昂贵不在于金钱方面，而在于空闲的关注时间太少。除了极少数被报道的事件人们记忆深刻以外，对于大部分事件而

言，能够形成共享记忆的只是个别情况，被遗忘才是常态。

当然，一些外部记忆并非广泛地散布在出版物或者广播中，而是作为个人的一种替代性记忆，帮助他们，其他少数人记住一个人写下的日记，写给一位朋友的书信，一家公司的会计账簿，或者是祖母的秘方。直到18世纪末，这种外部记忆在大多数地方仍然还非常昂贵，因此只有上层社会才拥有。

这一局面的形成，是否识字是其中的原因之一；纸张高昂的价格则是另一个原因。工业化的纸浆厂降低了纸张的价格，因此也使这种记忆变得更加便宜。但是随之而来的还有一个后果严重的隐藏的缺陷。纸浆厂生产的纸张不防酸，而且时间久了这种纸的破损也相对较快。写在这种纸上的文件，具有一种天生的自毁机制和一个自动的有效期限。这也解释了为什么那时候写在廉价纸张上的文档，今天保存下来的如此之少。因此，虽然19世纪比之前的年代产生了更多的外部记忆，但是由于多种条件的限制，长期的、能够承受得起的记忆仍然无从知晓。

介质记忆：摄影、磁带与胶片

19世纪和20世纪发明的模拟存储器（相对数字设备而言），对于改变这一局面作用甚微。

我父亲生长在20世纪30年代的欧洲。当他到了十岁的时候，我的祖父，一名巡回法院法官，送给他一台柯达布朗尼相机。我祖父教给我父亲的第一件事就是，在拍照的时候要非常慎重，因为每拍一张照片都得花钱。他认为，胶卷应当留着捕捉最令人印象深刻的时刻。在后来的几年里，当时十几岁的父亲大概只拍了36张照片——他拍下了重要的家庭事件，以及他爬过的山。大约四十年后，我得到了我的第一台相机，而我也被这样告诫，每张胶卷都是很贵的，我应当节省地非常小心地去拍照。然而，我拍照时远远没有小时候的父亲那么慎重，所以我青少年时期那十年的照片塞满了一个厚厚的相册。

今天的青少年没有这么多的约束——他们数码相机中的记忆卡能存储几千张照片，而且用起来也很方便。但是在20世纪70年代的模拟世界，创建长期的记忆依然很昂贵。

快速地回顾一下外部记忆设备的历史就能发现，我的经历属于一般情况，而不是例外。在19世纪中叶，肖像摄影取代了大部分的肖像绘画；随着照相馆数量的激增，照相成为人们更能够承受、更容易接触到的事情，在需求比较集中的城市地区尤其如此。随着照相的兴起，最早的印象派画家只好试图去捕捉观看者的感觉，而非某个人物或风景的清晰外观（因为这是相机的优势）。但是并非每个人都赞赏捕捉视觉记忆

的普及。

法国诗人夏尔·波德莱尔写道，“我们肮脏的社会已急不可耐，像自恋的那喀索斯附体一般，急不可耐地对着金属碎片上自己琐碎的影像顾影自怜。”但是当时采用的摄影技术银版照相法也并不便宜。这些“耐久的纪念品”，正如它们在18世纪50年代登的广告那样，当时价格在50美分到10美元之间，相当于今天的100美元到2000美元（按2006年的货币换算）。这在当时需要花费一名英国工人一个星期的薪水。此外，由于当时还没有发明底片，银版照相法拍摄的照片是独一无二的，因此也就无法复制。几十年之后，福克斯·塔尔伯特的纸摄影法采用底片，取代了银版照相法。这才使伊士曼·柯达在1888年发明了普通版照相机，所以人们才能够使用相机创建他们自己的视觉记忆。之后在1901年，又出现了大受欢迎且极其便宜的柯达布朗尼箱式相机。人们最终拥有了自己能够支配的技术手段，以极高的清晰度创建视觉的外部记忆。但那时流行的摄影仍然是黑白的，这一局面维持了几十年。

1900年左右摄影底片的价格是4美分一张（换算成2006年的价格，大概是5美元）这对于记录一个特殊的时刻而言，这个价格的确还算合理，但也不便宜。在一个社交聚会上拍十张照片并且洗印出来，按今天的价钱算大概需要花费100美元，而且这还不包括相机本身的花费。到了20世纪40年代，黑白照片的价格，按今天的价钱算是一美元一张：可以承受得起，但当然还是需要花钱的。黑白摄影的历史与彩色摄影的历史非常相似。

一般人在20世纪30年代开始接触到电影摄影机，但是那时的摄影机非常昂贵。二十年后，初级版的摄影机和投影仪的价格由于大量生产都降到了200美元左右（按今天的价钱换算），当然这还不包括最昂贵的部分——胶片。甚至到了20世纪70年代，当业余录像制作达到顶峰的时候，胶片依然很昂贵。一个小时长的录像（当然是没有声音的）在美国的许多地方都要花费大约300美元。在这么高的一个价格上，人们不可能像几十年前使用相机抓拍特别事件与非凡时刻那样，在他们使用电影摄影机和胶片捕捉什么事物之前肯定会三思而行。

磁带录音机的出现，让20世纪50年代人们才开始接触到这件新鲜事——将声音记录到磁带上，但就像使用工具去捕捉视觉记忆一样，记录声音在最初也非常昂贵。到了60年代，标准的磁带格式——如八轨式与紧密卡式，才使业余爱好者记录声音变得既普遍又可以承受。到了70年代，磁带仍然需要两三美元一盒，而且空白的磁带也是需要花钱的。这种状况一直持续着，直到90年代末模拟音频时代结束为止。

模拟视频，与图像和声音相结合，遵循了当下常见的模式。最初它

极其昂贵：在20世纪80年代初期，摄影机的价格按今天的货币换算大约需要5000美元。这种设备的价格随着时间在下降，直到20世纪90年代中期进入了大众消费者市场。电视机，如今这一家家必备的收看设备，随后在家庭中普及开来。磁带的价格也变得可以承受了。人们开始用磁带录下一个小时又一个小时的事件，从特殊的事件到微不足道的事情。

然而，随着捕捉记忆的设备与媒体的价格的下降，以及能够捕捉的记忆容量的激增，之前那些隐藏着的代价开始显现出来。有了几十万张的照片与大量的磁带录音，以及随后出现的大量录像带，有效地提取记忆成了一个问题。如果用户想要提取他们制作的外部记忆，他们必须花费时间将照片贴在相册里，或者至少将它们整齐地放在信封或者盒子里。对于哪些内容存储在哪些磁带中，他们必须保留准确的记录，并且每当作了任何改动之后都需要更新记录。这绝非小事。

很多畅销的录音带播放器内置的计数器众所周知都是不精确的。相同的道理，随着录像带的录像能力从几分钟上升到几个小时，忘记哪些内容确切地存储在哪里的可能性也在增加。因为磁带是有序媒介，在一卷磁带中找到某个特定的位置不仅需要详细的记录，而且需要耐心，毕竟倒带与播放都需要耗费时间。

毫无疑问的是，人们捕捉并记录在各种类型外部记忆设备中的信息量，在最近的四分之一个世纪中有了大幅度的增加，但是在模拟时代，有效的记忆既复杂又耗时，因此也很昂贵。记忆仍然比遗忘要困难许多。

自人类早期开始，我们就已经尝试记住并保存我们的知识，将其存入我们的记忆中，同时我们也已经设计了众多的设备与装置来帮助我们记忆。然而，千年以来，遗忘仍然比记忆更简单，成本也更低。

取舍之道

随着新出现的帮助我们记忆的工具与设备的不同，我们记住了多少又忘记了多少也一直在改变。但是，从根本上讲，我们记住的是那些足够重要，以至于我们愿意花费额外努力去记住的信息，而其余大部分的信息都被遗忘了。随着时间的发展，记忆渐渐只比遗忘困难一点点的事实，却帮助我们人类避免了一个根本问题，那就是：如果我们能够记住的话，我们是否将要永远地记住所有的信息？但如今，我们已无法再避免这一问题了。

---

[1] 吉尔伽美什是传说中的苏美尔国王。——译者注

[2] 因维苏威火山大喷发而埋没的古城。——译者注

[3] 因一次突发的火山喷发而被火山灰完全埋没。——译者注

[4] 据《中国出版史》记载，中国毕昇早在11世纪40年代就发明了泥活字印刷，远远早于古登堡15世纪50年代发明的铅活字。——编者注

第二部分 大数据时代如何进行信息的取舍

03 世界已经被设置成记忆模式：数字化记忆发展的4大驱动力

人类对完整记忆的需求一直在持续上升，这让如今的世界已经被设置为记忆模式。海量的数字化记忆不仅唾手可得，甚至比选择性删除所耗费的时间成本更低。这是一个几乎失去了遗忘动机的时代……

没有人可以不交流。在数字时代，只要在沟通过程中运用了数字设备，交流产生的信息便会被添加到我们的“数字人生”中。

如今，对于成千上万的用户而言，将数字信息重新组合都变得触手可及了。但是这些新文化产品的创造者，将信息与其原有背景剥离开，让它与一个全新的背景相联系。

完美的信息控制即便在模拟时代也只是一种幻想，除了对于最坚定的隐士。而在数字时代，作为个人，我们控制信息的能力再次被大大减弱了。

小黑盒子与麦克斯存储扩展器

小型计算机之父戈登·贝尔（Gordon Bell）被称为“计算机界的弗兰克·劳埃德·赖特[1]”。他引领了美国数字设备公司（DEC）的传奇产品——PDP（程控数据处理机）与VAX（一种支持机器语言和虚拟地址的32位小型计算机）系列的发展；塑造了美国国家科学基金会（National Science Foundation）在跨越整个美国的网络计算机方面的工作——这就是将要成为互联网的东西。

在20世纪90年代中期，当达到退休年龄时，贝尔加入了微软公司的研究部。他是非常睿智的长者和工程师，常常面露可爱的微笑，有着无穷的精力，伯父般的慈祥，以及深邃的洞见。唯一奇怪的是，他的脖子上总是挂着一个烟盒大小的小黑盒子，这是戈登·贝尔最明显的标志。在他70岁的时候，这个小黑盒子再次肩负了一个使命，那就是：记住他生命中的所有片段。贝尔的目标无异于消灭遗忘。他说，“我相信，这就是对个人计算机终极意义的探索，那就是记录一个人完整的一生。”

近十年来，贝尔尽可能将他所能接触到的所有信息都存储在他的电脑中。这是他与微软研究部的同事合作的长期研究项目“我的生活片段”（MyLifeBits）的一部分。贝尔扫描了几乎所有的纸质笔记与手册，包括800页他的个人健康记录，并将它们存储到硬盘里。他保存了他发出的或收到的所有电子邮件的副本（超过12万封），以及他访问过的每一个网页的镜像。他以录音形式记录，并以数字形式存储了很多他与其

他人之间的对话；而且他戴在脖子上的小黑盒子实际上是微软开发的数字摄像机，这个摄像机每30秒会拍下一张快照，或者每当有人接近贝尔的时候也会拍照。就这样日复一日地拍照，他估计自己文件夹里的这类照片已经超过10万张了。

机器充当人类完美记忆“假肢”的思想并不新颖。在20世纪30年代晚期，英国科幻小说家赫伯特·乔治·威尔斯（H.G.Wells）描述过一个“世界脑”（world brain）的概念，通过它，“整个人类的记忆能够被每个人获取到”。几年之后，第二次世界大战期间一位出身名门的科学管理者范内瓦·布什（Vannevar Bush）通过论证，对今后可能出现的完美记忆机器进行了最有影响力的描述。在一篇1945年刊载在美国《大西洋月刊》上的文章《诚如我们所思》（*As We May Think*）[2]中，布什描述了一种他称为“麦克斯存储扩展器”（memex）的机器，它“使人类能够接触并控制每个时代传承下来的知识”。用户将不仅能够查询麦克斯存储扩展器，还能够不断地将信息添加到它的存储器中。与威尔斯的“世界脑”非常相似，对于布什而言，麦克斯存储扩展器将“实现一种方法，通过这种方法人们能够制作、存储并查询人类的记录”。

然而，范内瓦·布什的麦克斯存储扩展器从未被制造出来。技术的方向不在那里，而且布什也与其他冒险者一起，精力不再集中于此。但是半个多世纪之后，戈登·贝尔实现了布什的大部分预想，而且走得更远。他成功地制造了一个存储他生活信息的巨大外部存储器，他可以随时在几秒钟之内访问这些信息，仅需要按几次键盘就能实现。那么，在布什与贝尔之间的半个世纪中，什么东西改变了呢？

我们已经从模拟信号时代进入了数字时代。现代技术已经从根本上改变了能够被记住的信息的内容，改变了记住信息的方式，也改变了记住信息所需要付出的代价。随着经济方面约束的消失，人类已经开始大量增加他们存储到数字外部存储器中的信息量。

取舍之道

便于提取的数字化记忆数量，在过去十年间有了大幅的增加，每年的增长率持续高涨。2002年的一项研究估计，每年大约有5艾字节[3]信息被存储，年增长率达到了令人错愕的30%左右，这些巨大的信息堆中超过90%是数字形式的。如果这个数据是相对准确的估计，那么，仅2005年一年产生的信息量，就应当已经达到了10艾字节。这仅仅只是我们每年新添加到全球数字化记忆中的信息量。很明显，记忆已经成了常态，而遗忘反而成了例外。四种主要的技术驱动力促进了这一转变，它们分别是：数字化、廉价的存储器、易于提取以及全球性覆盖。

驱动力1：数字化

将我们的时代称作“数字时代”似乎已经是老生常谈了。从模拟信号到数字信息的转变，即便没有影响到经济成分和我们生活中的所有方面，也已经影响了其中的大部分。

在40年前，数字化还只是一个相对晦涩的话题。而今天，如果主流文化没有数字化，几乎不可想象。我们听的是数字化音乐，拍照与录制视频使用的也是数字设备。我们创造的大部分文件，从简单的信件到复杂的设计，都是以数字形式存储的，并通过数字网络传送给收件人；全球电话网络也是基于数字代码；电影正在走向数字化；电视与广播也是同样如此。医疗诊断，从磁共振成像（MRI）到测量我们的血压或者体温，都以数字形式呈现其结果。甚至是孩子们玩的无线电操控的模型飞机、玩具火车与有轨电动模型赛车使用的也都是数字技术。

麻省理工学院媒体实验室创始人尼古拉斯·尼葛洛庞帝（Nicholas Negroponte）将他的畅销书命名为《数字化生存》（*Being Digital*）也就不足为奇了。尽管数字化模式现在已经是主流，但是数字化模式本身以及为什么它与模拟信号不同，仍然像理解记忆状态急速上升的重要性一样复杂。

在模拟信号的世界，存储的信息往往是对原始信号相对准确的复制。比如，为了保存声音，它被转译成电信号，电信号随后又被用于磁化磁带的表面。不幸的是，当模拟信息（即以模拟信号的形式存储的信息）在随后被处理、存储与传送的时候，随机的变动，确切地说是噪声，总是被加进来，最终致使原始的信息无法理解。

我们当中那些对于模拟音乐录音带有印象的人对此太了解了：对一盒真正好的音乐磁带不断拷贝进行复制产生的，不是另一盒同样拥有良好音质的磁带，而是一盒音质不太好的磁带。在每一次复制的过程中，噪声慢慢混入录音中，降低了它的保真度。

复制模拟录像带也会导致类似地问题，就像被用来复印的文件本身就是一份复印本会遇到的问题一样。这就好像在模拟信号的世界中，为了有意记住而制作的每一分拷贝都无法逃脱一部分的遗忘——造成的后果就是，每复制一次就会对内容产生一次轻微而随机的破坏。重要的是，回溯到首份拷贝（而不是对拷贝进行复制）并不能消除这个问题，因为复制不仅将噪声混入到了拷贝中，也将噪声（尽管程度很低）混入了原始版本中。这也就能解释当一个人播放一盒模拟录音带太多次之后发生的情形：存储的模拟信号越来越弱，并且导致声音的质量也受损。所有这些影响都能够通过设计完好（而通常也很昂贵）的设备得到减

弱。

一套高端唱片播放器与放大器，包含一个设计完好的扬声器，播放的是需额外支付费用的高端唱片。在模拟信号时代，这样一套高端设备可以产生优质的声音，但是这种模拟记忆需要承担巨大的代价——包括以设备的形式，以及由于随着时间的推移，唱片的音质也逐渐在衰减。同时，由于随机的变动无法从原始信号中区分出来，这种“噪声”永远无法完全被消除。

数字信息则完全不同：噪声能够被避免，所以质量不会随着时间衰减。其中原因也很容易看清：由于信息是数字化的，所以它可以近似为它的各个部分组成的离散集合。

就音频来说，通过在每秒钟内对频率与振幅测量上千次，声音被数字化了。这些测量中的每一次测量都是一个“样本”，表示某个特定（而且也非常短暂）时刻的声音。将众多的样本以正确的顺序合在一起，就复制出了一种近似原始声音的副本。如果每秒钟的抽样数足够多，对于人类的耳朵而言，复制的声音与原始的模拟声音根本无法区分。

类似地，一幅图像，能够通过使用一个由多个或黑或白的小领域（俗称像素）组成的栅格（比如，1000×1000）被数字化。这就能创造一幅原始图像的黑白近似图像。如果有很多这样的像素，而且每个像素都足够小，那么作为结果产生的图像将会非常好。传统黑白报纸上的照片的原理与这种方式有几分相似——当一个人近距离地看报纸上的照片时，一个个小圆点是清晰可见的。

这种将感觉信息（模拟信号）转译成离散状态的方法，就像一个既可以是黑色也可以是白色的像素一样，最初听起来可能感觉比较繁琐，但是它却提供了一个巨大的优势。它极大地改善了“噪声”的问题，在印刷领域我们称之为“噪点”，也就是随机变动的问题。

以一幅数字图像为例，它由1000×1000个或黑或白的像素点组成。当把它在一台激光打印机上打印出来的时候，我们就把它转回到了模拟形式，也因此引入了噪点。打印机可能不会在所有像素点上都完美地工作，而且其中一些黑点可能不会像另一些黑点那样黑。如果我们再次对这张打印出来的图像进行扫描（把它数字化），我们能够告诉扫描仪一个由单个像素点代表的信息位（information bit），它只能是黑色或者白色，而不会是灰白色或者淡灰色。因此，当我们的扫描仪遇到一个不尽完美的、打印出来的灰色点时，它能够自行将那个点的信息更正为原来的黑色，这样噪点就被消除了。

有了数字信息，大部分随机变动，即噪声或噪点，就能够被过滤掉，就像任何信号（如“灰色”）与某个特定值（如“黑色”）足够接近

时，就能够被理解成那个值。

如果这些特定的预先定义的值或者离散状态尽可能少的话，识别噪点的效果最好。假设数字化系统根据暗度将一个像素编码为白色，灰色或者黑色。噪点可以改变其信息，以至于一个白色的像素可以变成灰色像素或者黑色像素，一个灰色像素可以变成白色或者黑色像素，或者一个黑色的像素变成白色或者灰色像素。

这样，一共存在六种由噪点造成的可能错误状态。我们将这种情况与只拥有两种状态（黑与白）的系统相比较。在后一种系统中，噪点可能将一个白色像素变成黑色像素，或者将一个黑色像素变成白色像素。它只有两种可能的错误状态。数字化系统无法告诉我们某个特定系统有多少个预先定义的状态。它可以是100个，10个，或者5个，但是至少必须要有两个状态。仅基于两个离散状态进行信息编码采用的是二进制系统，比如黑与白，开与关。因为它使用的离散状态数最少，所以它能更好地避免噪点。这也是为什么今天所有常见的数字化系统都使用二进制代码，以及为什么使用二进制系统数字化的信息已经成为与数字化信息这一术语同义的众多原因之一。

剩余噪声的问题可以使用复杂的“校验和”[4]与类似地方法进一步降低。结果是，为了实际运用数字化信息，我们对噪声的修正也习以为常。它的影响是深远的：这与在模拟信号的世界不同，第一份数字拷贝是一份精确的复制品；每个比特（二进制的信息单位）都是原始信息的精确复制。几百次复制数字原始版本拷贝的拷贝，结果产生的副本仍然与原始版本一样完美。质量没有下降，而且复制并不附带对原始版本的损害。提取信息也不具有损害性：我能够百万次地听相同的数字音乐文件，而且也不会磨损；虽然我的播放设备与耳朵可能会随着时间变化受损，但是包含信息的文件却不会受损。

乍一看，数字化的这种明显优势还伴随着一个明显的缺点。毕竟，数字化创造的仅是原始版本的一个近似版本。至少在原则上，它不如模拟方法准确。但是，这种缺陷的程度取决于与原始版本的相似程度有多好或多坏。如果声音被抽样的频率只是每秒钟几次而已，那么，作为结果产生的数字录音对于人类的耳朵而言，就几乎不可能被识别，而且极大地落后于模拟方法。另一方面，如果声音被抽样的频率是每秒钟成千上万次，那么它就提供了这样一幅数字景象，当听到这个声音时，它提供了一种极好的音频体验。

如果由100×100个只能是黑或白的像素组成的栅格产生一幅图像的话，大量的信息将会丢失，同时作为结果产生的数字化图像将会是原始

版本的一个极其简陋（而且无法令人满意）的近似版本。但是，假设一部数码相机拍摄一幅图像时使用了4000×4000个像素的栅格，每个像素点代表几百种离散颜色状态中的一种，而非只是黑色或白色的话，那么效果就会好很多很多。这样的一幅数字图像，至少对于未经训练的人眼而言，是无法将其与原始版本进行区分的。

利用更多的信息，即以更高的频率对声音进行抽样，或者使用具有更高像素的传感器去拍摄照片，是解决近似这一障碍的关键。这也是为什么音乐CD（以每秒钟44100次的频率对音乐进行抽样）与所有设备（但除了最精密昂贵的设备）上播放的模拟磁带与录音相比，能够提供了更好音质的原因。对于使用千万以上像素去数字化一张图像的数码相机而言，同样也是如此。

类似地，与常规模拟信号电视提供的区区13万像素相比，最新一代的便携式摄像机现在能够以前所未有的“高清晰”质量拍摄，每张图像大约200万个像素。当然，这种数字化系统的保真度不仅取决于抽样数，还取决于声音传感器与图像传感器以何种程度将声波与光线转化成数字信息。这也是为什么对于数码相机而言，更高的像素并不必然能转化为更高图像质量的原因——图像质量还取决于使用的数字化传感器的质量。

由于没有噪声的问题，所以数字信息优于模拟信息。只要技术可以足够快速地将模拟信息转化成数字信息（反之亦然），并且存储的数字信息量能够满足原始（模拟）信息的充分近似的需要的话，数字化系统注定要取代模拟系统。一旦模拟声波能够被处理成几百万个数字样本，存储在一个负担得起的介质（比如一张音乐CD）中，并且被一台能够将数字信息转回模拟声波的设备（CD播放器）所播放，那么整个音乐产业都将转化为数字形式。我们已经在文本处理、摄影与视频领域看到了类似地转变。

但是，数字化系统的优点远远不止如此。在模拟信号时代，处理、存储与提取过程因不同类型的信息而不同。这就导致了关于不同信息类型的专用设备与一个专门的市场生态系统的发展。

消费者使用录音机与空白磁带存储和提取声音。预先录制好的磁带由音乐出版商生产，并通过专门的音乐商店出售。消费者使用专门的相机与模拟胶卷拍摄照片，这是当时照相冲印实验室开发出来的。人们通过钢笔或者打字机将墨水印在纸上来创建文件，而纸张、墨水、钢笔与打字机都是通过专门的“办公”用品商店出售的。消费者看的录像是通过模拟便携式摄像机录制在模拟磁带上的，随后再显示在模拟电视机上——所有这些设备都需要从电子商店里购买。

这种根据特定类型、煞费苦心的信息处理的方式是有效的，但是它的效率显然比较低：我们在家里和上班的时候需要使用很多种不同处理、存储与提取信息的设备，而且它们中的每一个都需要正确的操作与适当的维护才能运行。如果我们还剩下一盒备用录像带，我们也不能用它来记录文本，就像我们无法用纸张去存储音乐一样。

在数字信息的世界，所有信息都以二进制信号存储，而且所有设备，至少原则上，都能够处理这些数字信息。只要信息能够被数字化，它就能够被存储在同一台数字存储设备上，而不管信息是声音、视频、文本，还是其他类型。人们可以将贝多芬的音乐与他们的婚礼录像，以及他们最新的商业计划都存储在同一个硬盘上，而无需担心这些信息有混合或者被损坏的风险。

标准化通过大规模生产，为数字存储设备制造商提供了巨大的市场机会，因为这些设备能够被用于存储所有类型的数字信息，这进而产生了效率。由于生产的数量非常巨大，这些设备的单价已经有了大幅的下降。标准化还推动了对一种多用途信息处理设备的需求，只要连接上了必要的输入和输出设备（麦克风与扬声器，相机与屏幕，键盘与打印机），就能够编辑文本、图像、声音与视频。个人电脑数量急剧上升，并占据信息处理领域当前的主导地位，至少部分受益于数字化驱动的这种对通用设备的需求。

标准化的好处，还延伸到了通过或大或小的网络进行信息共享与分配的领域。模拟信息在相互分离的基础设施中，依据不同的信息类型进行共享与分配。报纸通过卡车送往各地的销售员，电影用船运到各家电影院；模拟广播则通过发射器系统来传播。相比而言，数字信息不管什么内容，都可以通过相同的数字网络进行传输。共享一个传输网络而非采用单独的分配网络，明显更为高效。

数字信息的标准化还有另一个好处：与模拟信息相比，它更能经受住“未来的考验”。只要一个信号能够被数字化，它就能够用诸如个人电脑这样的标准化数字信息设备进行处理，用诸如硬盘这样设备进行存储，并且通过互联网在世界范围内进行传输。

互联网本身的演变也是一个非常好的例子。最初，互联网被用于多个（远距离）用户共享一台电脑的信息处理能力。随后，在同一个网络中，文件传输与电子邮件被添加为新的服务。而当蒂姆·伯纳斯-李（Tim Berners-Lee）发明出万维网的时候，它依靠的就是既有的互联网基础设施，就像后来的音频流与视频流，以及网络电话（即IP电话）一样。当然，随着时间的推移，互联网的传输管道也已经更新修改了，但是它的基本建设原则仍然保持不变。

互联网作为一个全球数字网络，能够经受住“未来考验”的基础是，它不会被弃置，不会被取代，甚至不会从根本上被改变。这样就能够使即时通讯、在线游戏，甚至在诸如《第二人生》（*Second Life*）游戏这样的复杂虚拟世界中畅游成为可能。

消除噪声同时保持保真度，并通过标准化、大规模生产以及更好的资源分配使得效率得到大幅上升，数字化系统已经使得从根本上改善信息处理、存储与提取成为可能。但是，它的影响超越了在经济学上的显著作用。通过消除噪声，模拟技术不断增加的固有遗忘，原始版本与拷贝的概念，都快速地成为过时的概念。所有的数字拷贝都无法与原始版本进行区分。这为每个想使所有的信息，包括“拷贝”与“原始版本”接近完美的人提供了好处。数字拷贝的完美性极大地方便了高质量信息的分配。多亏了它，今天我们才能够享受到高保真度的音乐与高清晰度的视频。我们能够负担得起的相机，可以以模拟时代的业余爱好者几乎不可能接触到的分辨率去拍摄照片。而且我们能够与成千上万人分享这些照片，而双方都无需忍受由广泛传播与反复复制造成的质量下降。

并非每个人都为数字信息无损、廉价与易于复制的特性而高兴。内容的拥有者担心人们复制了内容而不付钱。在模拟时代，噪声，这种缓慢但不可避免的遗忘，充当了大规模非法复制的障碍。试想一下，一盒非法模拟音乐磁带的拷贝由于模拟技术对音质的降低，听起来效果平平而不再具有高品质；这可能会促使消费者去购买一份直接源自母带的正版拷贝（或者更可能是从母带上移植的拷贝）。另外，如果一份非法的拷贝能够与原始版本一样好，那么让消费者去购买正版就要难得多了。要做到这一点可能需要在正版中提供额外的东西（大概出版商已经通过苹果公司的iTunes音乐商店以及类似地合资企业进行了实验），或者直接叫警察了。

然而，不考虑这些版权的问题，数字信息易于在世界范围内获取与分享的特性极大地增加了我称之为“共享的社会记忆”的东西。社会记忆不仅在规模上增加了，而且还成为了全球的共享记忆。

几年前，我在文莱首都闲逛时，经过那里唯一的一个大型购物商场。我注意到，一家商铺的广告醒目地告知消费者，他们在私售美国的一部著名电视剧《欲望都市》。起初，我对于这么露骨的电视内容能够在一个传统的伊斯兰社会如此公开地出售大吃一惊。但是，如果考虑到如今全球信息都能通过卫星电视、私售的DVD与CD，以及互联网进行流通，就很容易理解为什么文莱的人们也想跟随《欲望都市》的女主角凯莉·布拉德肖的视角，穿过纽约曼哈顿、美国辛辛那提，或者罗马的约会场所了。

取舍之道

由于大众媒体的全球性覆盖，共享记忆即便没有全球化，也早已经国际化了。而且数字化系统已经使一个更大、更为全球化的共享记忆成为可能，同时，这种可能性远大于模拟时代的可能性。另外，它可能会奇怪地助长了这种共享记忆的碎片化，因为人们可以去选择，将自己在多大程度上暴露在这些社会记忆的哪一部分面前。

数字拷贝与原始版本一样好的特性，也改变了人们感知信息的方式，尤其是对那些成长于数字时代的人而言。如果在模拟时代，拥有很多书籍或者音乐磁带或者电影录像带会被认为很酷，那么在数字时代，能够在它们的基础上去构建新东西才酷——将我们信息文化的一些手工产品结合成一些新的或原创的东西。史蒂夫·乔布斯曾对音乐有过一句精辟的描述："扒歌，混制，刻录。"（Rip，mix，burn.）也许还可以加上每个人自己的内容。尽管强调的重点放在了混合（mixing）、重新组合（recombining），以及著名的施乐公司帕洛阿尔托研究中心（PARC）前任领导人约翰·希利·布朗（John Seely Brown）所提出的再创造[5]上，这些过程的价值其实来源于其各个部分的组合与重新组合，而不必然来自各个部分本身。希瓦·维德海纳森（Siva Vaidhyanathan）提供了"妖精编辑"（Goblin edits）的例子，富有创造力的俄罗斯艺术家将西方最卖座的电影重新用俄罗斯语配音，赋予了这些电影全新的故事。随着数字化文化强调重新组合与共享而非拥有，人们正在利用数字化系统带给他们的信息权力。

总之，数字化系统是新一代信息处理、存储与提取的促成者，而且共享这些信息极大地超越了模拟时代信息的共享。

驱动力2：廉价的存储器

在20世纪40年代早期，当数字处理技术刚刚开始起步的时候，数字存储的代价极其昂贵。第一台成功的商用计算机UNIVAC（通用自动计算机）的主存储器有12000个字符（字节），同时一盒磁带驱动着它的大容量存储器。一盒磁带能够容纳一兆字节（megabyte）。当时UNIVAC计算机的售价是100万美元，换算成美国现在的价格大约相当于750万美元。

磁带是一项很有用的创新，但是除了其价格昂贵之外，它们提取数据的速度也很慢：如果某一块信息被存储在一盒1200英尺长的磁带的末尾，那么，想要提取这块信息就要把整个磁带都放完。工程师们便开始寻找能够提供更快存取时间的计算机存储技术。

磁鼓（magnetic drum），是这个寻找过程的第一步，但是它不仅昂

贵而且存储容量有限。1957年，IBM推出了305计算机，这款计算机以磁盘作为存储设备，能够提供高达5M的空间，价格在100万美元左右（按2006年的货币换算）。在20世纪50年代，仅存储器的费用就高达大约每兆七万美元；到了1980年，这个价格已经降到了500美元以下（所有价格均换算为2006年的货币），连25年前价格的10%都不到。20年后，到了2000年，存储器的价格已经骤降到1美分左右，是1980年价格的五万分之一。到了2008年，一兆的存储器的价格已经降到了0.01美分。50年间，存储器的价格每两年大约就会下降一半，而存储密度则增加了5000万倍，从1956年的每平方英寸2000比特增加到2005年的1000亿比特。

硬盘的价格与存储器的大小不是成正比的。因为硬盘本质上是由一个电动机驱动的磁盘，并且以每分钟几千转的速度旋转，总成本的一部分往往花在这些必需的机械部件上。基于芯片的存储器则不具有这一缺点。这也是为什么我们看到，在过去10年间对基于不挥发硅的存储器，即所谓固态存储器的使用，存在一种稳定的增长。数码相机，手机与音乐播放器都采用这种“闪速存储器”。最初它相对昂贵，逐渐存储卡每兆的价格与硬盘价格一样有了显著的下降，降到了2008年大约每兆三分之一美分——这一价格与成本之比是硬盘2001年就达到的水平。然而，固态存储器具有一个优势，那就是它不需要机械部件。

自从硅基集成电路出现以来，工程师们已经能够缩小这些电路上的单个元件的大小，这样也就能够在相同的空间内集成更多的元件。1965年，一位名叫戈登·摩尔（Gordon Moore）的工程师写道，“自集成电路出现以来，它的复杂性大约每年会增加一倍。每个功能的集成电路的成本已经下降到当初的几千分之一。集成电路复杂性的增长率可以预计将在未来几年内发生变化，新的增长率可能大约是每两年增加一倍。”这个本被当作是推测的观察已经适用了40多年，而“摩尔定律”如今已经成了一个简称，代表着每18个月集成电路的复杂性就增加一倍，并且自有集成电路以来我们已经目睹了这个定律的正确性。

戈登·摩尔不仅说得准，而且做得也好：他与人联合创立了今天的芯片制造巨人英特尔公司。因为固态存储器没别的，就是一个复杂的集成电路，所以摩尔定律同样也继续适用于固态存储器。而当前技术唯一的根本缺陷就是，闪速存储器会缓慢地磨损。当信息已被删除或写入超过10万次之后它会变得不可靠。而对于一个更长期的存储媒介而言，这不是一个太可能发生的事件（因此也算不上什么缺点）。

几十年来，光学系统已在数字存储器中扮演了一个重要的补充者的角色。在消费者层面，这已经导致了CD（1982年被引入，拥有6.5亿比

特的存储空间）、DVD（1995年被引入；存储量是CD的7～13倍）和蓝光光盘（Blu-ray Disk，2006年被引入，比CD的存储量大大约40～80倍）的广泛使用。光学系统中存储能力增强的步伐相对缓慢并不必然是技术限制的标志，而是用户的偏好。光存储器的发展是基于可替换的媒介的（光盘）。当底层的技术从一个系统更换到另一个系统时，老式的可替换媒介记录的内容可能不能再播放，或者无法以新设备的用户所期望的质量播放。

这对于那些仍然拥有大量老式模拟VHS（家用录像系统）盒式录像带电影的收藏者来说，就是一种很常见的情况。由于消费者已经买了可录制的媒介，他们不愿去更换。这些可感受到的更换成本限制了这类存储系统的新一代产品能够多久一次获得市场成功。为了降低更换成本，设备制造商有意设计了DVD和蓝光驱动器，它们也能够读取较老的格式，因此至少保证了一些向后兼容性（backward compatibility），但是这些解决办法只是权宜之计。这种情况对于硬盘则截然不同。因为不存在可替换的媒介，当硬盘技术发生改变时，用户并不在意。对于他们而言，一个硬盘就是一个黑盒子。只要它能够存储和记住东西，底层的技术制造工艺是无关紧要的。因此，硬盘存储方面的新技术进步能够更快地被引入市场，而且无需惧怕更换成本。

直线下降的存储价格已经受到用户的欢迎。当价格已经降到私人需求也可以负担得起的时候，它就将长期的数字存储器带给了大众。而且，消费者也不太可能很快就用尽存储空间——等到一个人已经装满了她的硬盘时，又可以购买一个拥有两倍或三倍存储空间的新硬盘。戈登·贝尔认为，人们将不会再有硬盘空间不足的情况。贝尔可能会惊讶地看到，即便300G的空间被高清晰的视频填满是多么容易，但是在某种意义上他还是对的，那就是存储能力的增强正在继续满足人类对额外的存储空间的欲望。

在某些情况下，数字化存储信息变得比模拟存储更廉价。一页文本，单倍行距，包含大约2000比特的信息。激光打印一页文本的成本已经相当稳定，大约10美分，所以制造一兆字节文本的模拟版本大约需要花费50美元。1987年时，磁硬盘上的数字存储已经降到了每兆40美元，这已低于模拟存储的价格。从那时起，数字存储不仅更加便利了，而且比打印和保存传统的硬拷贝（hardcopy）更便宜。

以图像为例：打印一张常规照片大约需要花费30美分。1997年时，数字存储已经变得如此便宜，以至于存储一张数字格式的图像比常规的膜印刷（假如是一个存有3兆图像的文件）的价格还便宜。到2000年时，在硬盘上以数字格式存储一个音乐文件比存储在空白的模拟录音带

上还要便宜，这使得传统的盒式磁带录音机变得很不经济。而且到了2006年，在硬盘上存储消费者的视频已经比买空白的模拟录像带更便宜。

如今，数字存储对于消费者的使用来说，显然已经比模拟存储我们生活和思想的记忆更便宜。

这听起来可能令人很错愕，但是廉价的存储器还有一个更为重要的方面——存储空间。当硬盘制造商迅速意识到这一点并提供了一种选择时，消费者宁愿花与上次差不多的价钱买一个存储空间更大的硬盘，而不愿以不断下跌的价格再买一个与上次相同存储空间的设备。至少一旦存储器达到了一个消费者能轻松接受的价格水平时是如此。不是为了省钱，用户想的是利用数字存储去保存越来越多的信息。因此，硬盘的价格已经大致稳定了，但是我们使用的硬盘的存储能力则一直在提升。这就是为什么我们的个人电脑能够提供太字节（TB）级别，即100万兆字节（MB），的硬盘存储空间，而价格却差不多与20年前相同；我们的数字音乐播放器能容纳好几百小时的音乐；而我们数码相机的闪存卡里能够存储几千张高清的照片。

两个驱动因素使得这个不可思议的增长在廉价存储器上成为可能。

第一个因素在研发和生产方面。存储设备，不管是硬盘还是硅芯片，都是大规模生产的，而且仍然必须符合严格的生产标准。起初，生产容易产生误差，导致产量很低。随着制造商对生产工艺进行调整，产量提高了，而且这种规模经济拉低了价格。这一过程会一直继续，直到进一步细化生产工艺的代价变得比他们生产所获的收益还要高昂的时候。制造商还试图革新产品，以期在产量提高已经不可能的时候切换到新的产品上。这种新产品在存储能力方面能够提供明显的提升，但是最初，当生产产能提升时，生产的产量仍然很低。

另一个因素是，存储器市场的需求在增加。这使得市场推动制造商提升存储能力并降低价格。而且，因为存储价格下降导致的需求增加，使得信息存储和存储设备能够让新的人群对于负担得起——当一台个人电脑只需花几百美元时，就会有更多的人愿意购买。同时，不难获取到的存储空间和进一步的数字化，将使得将各种类型的信息（文本、图像、音频、视频和空间数据）转置到数字设备上成为可能，这又再次刺激了需求。比如，随着数码相机用户将照片存储在多用途硬盘上，数码相机的兴起已经提升了这一需求。

让我们简要回顾一下：数字存储已经变得如此便宜，以至于在数字存储器上存储信息，即便是全屏视频，比诸如纸张、胶片与磁带这类模拟存储信息的方式还便宜，这使得保存我们遇到和生成的海量信息成为

可能。就像戈登·贝尔，他在使用数字媒介捕捉自己的全部生活方面走得很远——其中包括几十年的记录在他笔记本和文件上的想法和笔记的数字版文件，十万张照片，几百个小时的音频，以及几千张他访问的网站的镜像图片。他估计他已经积累了大约100G的信息，这甚至都没有装满他那台普通笔记本电脑硬盘的三分之一。他可能还需要录下他每天的日常生活以获取更多数字信息。

对于我们而言，这种异常丰富的可获取的存储空间，使得我们行为上的关于外部记忆的常态很容易地从遗忘转换为记住。我们将我们工作所涉及的不同类型的文件保存到我们的硬盘上，因为某一天我们可能需要用到它们。存储信息已经变得极其便利，但是引导我们保存数字信息的还不仅仅是为了便利。事实上，存储信息的经济原理已经使遗忘变得异常昂贵。拿数码相机来讲：当你将相机连接到你的电脑上，把其中的照片上传到你的硬盘时，你通常会做一个选择：要么选择手动上传哪几张照片，要么选择让你的电脑自动地从你的相机里复制所有照片。后者或许可以使人稍微安心一点，因为你可以随后再仔细检查照片，并删除你不喜欢的，所以绝大多数人都会选择让电脑自动复制所有照片这个选项。从经济学上讲，这很容易理解。假设一个人查看一张照片并决定是否保存它只需要3秒钟，并且假设他将他自己的时薪等同于目前的平均工资，那么，仅仅花在决定保存哪些照片上的时间“成本”就超过了存储的成本（包括花1秒钟将照片复制到备份设备上）。

拥有这么丰富的廉价存储空间，使得连花时间去决定是记住还是遗忘都已经不经济了。遗忘——需要花三秒时间去选择，这对于使用数码相机的人们而言已经变得太昂贵了。

适用于个人存储的规律，同样适用于企业的存储。公司级别的数据库保存了数量极其庞大的可访问信息，其中有商业交易的信息也有外部和内部的沟通信息。电子邮件会被保存，而且扩展到移动通信设备公司也会存储即时信息，其中的一些信息实际上对于公司是毁灭性的。微软公司在针对它的反垄断诉讼期间发现了这一点，当时政府律师强制微软交出几百万封内部邮件。这种情形与大型公共组织中的情况颇为类似。

当比尔·克林顿将电子邮件引入白宫时，他万万没有想到几年以后，这成为了肯尼斯·斯塔尔（Kenneth Starr）无情调查了自己私人亲密活动的一部分，他的职员先前所发的一些电子邮件再次浮出水面。在一次有效撤销政策的过程中，小布什下令让他的白宫职员随意使用电子邮件，而他的行政系统后来承认电子记录已经（有人说是“轻易地”）被服务器删除了。其他政府工作人员，包括2008年共和党副总统候选人、阿拉斯加州州长萨拉·佩林（Sarah Palin）也曾通过电子邮件账户绕过保留记录

的要求。很多附加信息久而久之被丢失了，不是因为设备故障，而是因为软件不兼容，或者糟糕的数据管理。

据《纽约时报》报道，在美国国家航空航天局（NASA），40位被调查的高级官员中，有93%的人没有适当地保存他们的电子邮件，因此违反了联邦保留记录的要求。

取舍之道

数字存储器提供了一种廉价而可靠的存储方式，但是如果缺少适当的数据管理规程，包括经常并可靠地进行适当的备份，数字存储的好处就完全无法显现出来。而且正如专家所说，不像硬盘存储空间的成本那样，这些数据管理的成本不会每一年半到两年就减少一半。但是即便牢牢记住这些限制，很显然，不仅个人，就算像公共组织这类私密的机构也遭遇了永久而广泛的存储造成的后果。

在1996年到1997年间，微软的电子邮件Hotmail和雅虎的电子邮件首次登场，为任何人提供免费的电子邮箱账户。为了提供这项服务，两家公司开辟了几兆字节的免费磁盘空间，以存储即将出现的电子邮件。在几个月之内，有几百万人注册了电子邮箱。到2000年年初，图像共享服务开始提供类似地待遇，让用户免费地上传并分享数码照片。大部分这些靠广告提供资金的服务的标志，都是在中央服务器上提供免费的有限磁盘空间，这与下降的存储成本相一致。

2004年4月1日（不是愚人节的笑话），Google公司通过提供1G存储空间的免费电子邮箱账户震惊了竞争对手。这比仅次于他们的竞争者所提供的存储空间大几个数量级。当时1G的硬盘只花了Google几美元，另外还需要再加上数据管理、电力、备份和带宽这些成本。为了进一步加大筹码，Google宣布，在未来每天的每一分钟他们都将为用户增加存储空间。到2009年年初，每位用户的免费空间已经增大了7倍多，达到了7.3G，在Google的电子邮箱Gmail网页上有一个存储空间计算器。据Google称，他们的目标是“免费存储，让你永远无需删除邮件”。雅虎公司则通过为其用户提供无限的电子邮件存储空间予以反击。

当然，Google不仅仅只为用户存储电子邮件。他们还抓取网页，创造了一个巨大的可搜索的网页索引（Google就是这么成名的），并且还会在此过程中给网页拍下快照（以便用户在原始网页损坏的情况下还能看到“网页快照”副本）。此外，正如我前面所提到的，Google保存了每一次搜索请求，以及用户点击进去的每一个搜索结果。这需要一个巨大的存储空间，据报道Google一天要处理10亿次搜索请求。为了实现这一壮举，像Google这样的多家公司都在巨大的服务器群里运行着成千上万的硬盘。虽然他们都对详细情况保密，但是据说Google至少管理着50万

台服务器，以及数量多达100万的硬盘，每一个硬盘都拥有超过100G的存储空间，这样总共能提供大约10万T的数据存储空间。

这是大到令人难以置信的存储空间，但是纯粹购买硬盘的成本，可能大约需要几亿美元，但实际上并不会那么惊人。直线下降的存储器价格是其中的一个原因，大批量地购买则是另一个原因。而很重要但却被忽视了的第三个原因是，据悉Google已经避开了更为可靠但也更加昂贵的硬盘，转而选择了成品的廉价的一般用户级别的驱动器。这很好理解；对超过10万个这样的硬盘的仔细研究表明，这些驱动器是高度可靠的——虽然不像它们的制造商在广告上宣称得那么可靠，但是仍然极少可能出现故障。一个硬盘在损坏之前，平均能正常工作大约30万个小时。考虑到一整年也不过9000个小时，这代表一个硬盘能够持续运行的时间超过33年。

一个成品的廉价硬盘和类似地备份磁盘，在同一年内同时坏掉的可能性低于千分之一——这是非常低的可能性。这些结果也意味着，全世界的消费者所使用的上亿个成品的硬盘不仅仅廉价，而且是惊人可靠的信息存储设备。

专家表示，廉价存储的趋势将会一直持续到21世纪的第二个十年。全世界的公司都在为下一代数字存储设备而积极工作着。2007年，日本硬盘制造商富士通公司（Fujitsu）宣称，他们已经通过在常规的硬磁盘上运用一种新的存储机制，获得了每平方英寸1T的硬盘存储密度（storage density）。另外一个研究组通过研究纳米技术，以克服一些未来可能阻碍硬磁盘存储密度增加的物理限制。随着制造商在推进以基于芯片的存储器作为基于硬盘的存储器的中期替代物，尤其是在诸如笔记本电脑这类便携式设备上推行，关于固态基存储器的研究也在急速进行。这种硅基存储器拥有一个额外的优点，那就是极低的能源消耗，因为无需电力，一个磁盘就能旋转每分钟5000次或者更高——这在能源价格高昂的时代是一个很大的优点。

2008年春天，惠普公司宣称，他们已经成功生产出了一种大家一直在寻找的电路元件，那就是记忆电阻器（memristor），大约15纳米大小，而且他们还将能够把它压缩到4纳米大小，这可能是目前大规模生产的电路元件大小的十分之一——这将使得固态存储器的空间增大10倍成为可能。同时，美国IBM公司宣称他们的赛道存储器（racetrack memory）技术，将更进一步，可能导致存储器的空间增大10到100倍。这些还只是众多以求增大硬盘和硅基存储器存储空间的研究思路中的其中三个。此外，所有大型存储器制造厂的工程师们都在积极地研究替代性的存储方法，包括全息方法（holographic）、生物方法和量子计算方

法。而可能的中期成果是，存储空间将继续翻番，而存储器的成本大约每18～24个月下降一半，这将留给我们丰富而廉价的数字存储空间。

驱动力3：易于提取

“记住”不只是将信息存入存储器中，它还包括日后能随意又轻松地提取那些信息的能力。作为人类，我们所有人都非常熟悉，从我们大脑的长时记忆中提取信息是多么大的挑战。外部的模拟记忆，比如书籍，承载了大量的信息，但是要在书中找到特定的某一页信息则既困难又费时。许多存储的信息所隐藏的价值仍然被困在纸堆里，不太可能被利用。虽然我们可能已经存储起来了，但在实际中却无法轻松提取的模拟信息，跟被遗忘了没有什么区别。相较而言，从数字存储器中提取信息则容易许多，便宜许多，也迅速许多：在搜索框里输入几个词，点击一下鼠标，几秒钟之内一系列相匹配的信息就被提取出来了，而且以整齐的列表格式呈现在你面前。这种毫无障碍的提取方式极大地提升了信息的价值。

当然，人类也一直在尝试使信息提取变得更容易更简单，但是他们过去一直面临着重大的障碍。以书面信息为例。从碑刻和卷轴过渡到装帧的书籍，它们共同帮助保存了信息，而且确实提高了可检索性，但是书籍并没有彻底改变信息提取的方式。类似地，图书馆帮助积累了信息，但是在查询信息方面，图书馆的作用就没那么大了。只有完全进入了第二个千年，当藏书切实可用的索引（可能最初是从广泛的组织，发展到细分条目，后来发展到章节、希伯来诗歌与《圣经》）变得常见的时候，图书管理员才能够根据书名和作者找到某一本书。

作为现代图书馆兴起的一部分，人们花了几个世纪的精制提炼才发展出标准化的书籍编目和上架技术。主题索引，以及书中所用词语按首字母在字母表中的顺序排列所得的列表（当时还没有标明页码），都是在13世纪被引入并迅速传播开来。它可以帮助一个人根据他想更深入了解的某个主题范围去选择合适的部分。但是为每本书编制这样一个索引很费时间，而且索引也并没有让读者在一本书中找到想看的某个段落而变得容易。

又过了几个世纪，页码，这个在一本书中精确找到特定信息的前提，在16世纪初期出现了，历史真正进入了印刷革命时代。只有当主题索引和页码被结合起来时，现代书籍的样子才开始出现，也奠定了在大型的文本语料库（corpus）中提取信息的基础。

商业和国家的兴起，使得轻松又准确地对书面文件进行归档及提取成为必要。归档系统（filing system）是模仿图书馆的两级系统（two-

stage system）设计的：文件被放在档案柜里，每份文件根据名称、日期，或者类似地标准进行分类。这样的系统使得及时提取信息成为可能，但是仍需要很高的日常费用：文件必须被仔细地存档，否则它们可能就永远地丢失了；而且归档系统的组织必须非常小心地进行设置，因为以后就很难再改动它了。

因此，精通如何对模拟信息（从文件到书籍）进行归档和提取，在信息工作者当中成了一项很受欢迎的才能，而且团体组织会使用大量的财力物力去制造并维护信息提取所必需的工具（索引和目录）。到了模拟时代的末期，信息提取已经取得了长足的进步。通过结合书籍目录和主题索引，信息能够从图书馆中被提取出来。但是那是一个单调乏味的任务，昂贵费时又不精确，因为主题目录仅仅包含每本书最重要的主题词，而且书籍的主题索引依赖于作者和编辑的准确性。此外，这样的提取系统需要非凡的资源，团体组织可能负担但是个人很难承担得起，或者至少无法获得与团体组织相同的水平。因此，我们收藏的大部分个人信息——草图和绘画作品，照片和自制的视频，录有个人信息的盒式录音带、日记本、笔记本、剪贴簿、官方文件、银行对账单，以及合同——我们都无法像公共图书馆整理他们的藏书那样去仔细地组织、分类并编录索引。作为替代，我们使用了众多不同的分类系统。

信息提取中存在这些挑战的根本原因是，大部分的模拟信息存储是连续组织的：没有额外的线索，一个人不得不去翻阅，或者甚至为了找到某个特定的部分可能需要读完一整本书；缺少明确定义的归档系统，一个人必须一份一份，一页一页地去浏览文件，一张张地去查看照片。录音带和录像带同样也是连续的。如果一个人想找到某个特定的场景，再看一遍（可能速度可以快一点）是唯一的选择。

缓解这些挑战的标准方法是创建并维护单独的提取工具，可以让我们更加精确地提取信息：目录可以帮我们找到想要的书或者文件；主题索引和页码帮我们可以找到一本书中相关的段落；在录像带外面注明它的话题和内容，与记着场景和时间点的列表一起，帮我们可以找到想看的场景。但是这些提取工具既不灵活又复杂，而且需要时常小心翼翼地去维护。

在数字时代，所有这些都改变了，虽然它们不是在一夜之间改变的。尽管软件具有可延展性（malleable），而且数字信息处理几乎没有内在固有的约束，但是人类在设计第一代数字系统时，还涉及现实世界中信息组织的隐喻（metaphor）。因此，早期为数字计算机设计的文件系统，通过记录哪个数字文件存储在哪个目录下，复制了图书馆图书目录的核心元素。同时，软件工程师们开发出方法，以便在数字文件内部

轻松地提取和组织信息。他们也受益于硬盘的兴起，硬盘不像大部分模拟存储设备，它们允许直接地提取非连续的信息。

到了20世纪60年代末期，IBM公司的英国工程师特德·科德（Ted Codd）构想出了他称之为关系数据库（relational database）的理论概念，其本质上就是，信息以表格的形式被组织，行组合的是一份文件中相关的信息，而列存储的是结构相似的数据（比如名称、日期和序列号等）。通过被称为关键字（key）的标识符（identifier），一份特定的档案可以被提取，但是不像模拟目录那样，这些关键字是灵活的。如果一个人想通过事先没有创建关键字的信息，以图书目录中的“出版日期”为例，去搜索和提取文件，他也可以随后再添加关键字。拥有了这样的关系数据库，图书馆中传统的模拟卡片目录（card catalogue）就可以被更为通用的电子版本所替代。书籍的记录不仅仅可以通过作者的名字或者标题提取，也可以通过名字或者标题的一部分片段去提取，同时还可以通过诸如出版商、主题关键词这类信息，或者甚至它们的组合去进行提取。

起初，昂贵的数字存储器被明智地用作首要的信息存储器。随着数字存储器价格的直线下降，越来越多的信息被数字化，这助长了对更好的信息提取工具的需求。

在20世纪70年代的美国，著名法律联机检索系统Lexis和Westlaw使得他们的用户，可以接触到拥有数以万计法院判决书全文的庞大数据库，但是这些判决书仅仅只能用一组数量有限的关键字去检索。然而，用户想要的是通过搜索判决书中用到的词语去寻找相关的判决书，而不仅仅是搜索案件名称、案卷编号、日期，以及被录入索引的几个主题词。解决办法是，让数据库中每份档案的每个词都可以检索。这样的全文搜索仍然要求输入准确的词语或条款，因此这也不是一种一定能成功找到想要信息的方法，但是它显然比仅预设少量几个检索关键字的搜索方法要更简单也更强大。

最初，全文索引和检索仅被大型的信息数据库提供商所使用，但是到了21世纪的开端，它已经成为所有主要个人电脑操作系统的一个标准功能，为人们的笔记本电脑带来了强大的精确信息检索功能。同时，数字设备变得或多或少能够自动地生成并管理文件的元数据[6]，而不仅仅只拥有像名称或创建日期这类众所周知的类别。

数码相机能够自动为每一张照片存储一长列的元数据，比如相机光圈、快门速度、焦距长度、日期和拍摄时间。当照片文件被复制到硬盘上时，这些数据被保存了下来，并且以后能够被用于直接搜索和提取照

片文件。用户也可以自己添加元数据。越来越多的数码相机（尤其是手机上的照相机）为拍摄的每张照片添加了地理位置功能——经度和纬度，并存入元数据的序列中，以帮助我们人类日后辨认出这张照片是在哪里拍摄的。

最新的进展是，索引、检索和提取这些现代元素的无缝融合，这消除了文件水平上的搜索（类似于在图书目录中搜索一本书）与文件内部搜索（在一本书中搜索某个段落）之间有史以来的差别。这个融合的数字信息检索通过一个简单的整合界面进行运作；信息在所有水平上都是可以检索和提取的，而且几乎立刻以一种视觉上一致的方式被保存起来，而不管是哪种信息格式。这就是我们所熟悉的很简单的检索方式，在笔记本电脑上如此，无论是微软的搜索服务，还是苹果公司的聚光灯搜索服务。在互联网的搜索引擎中也是如此，比如Google公司简洁漂亮又功能强大的搜索界面。

取舍之道

这与模拟时代的差异是惊人的：通过便捷的提取，大量的信息，将不再是无尽的比特之海，也不会再有被信息淹没的风险，而是我们人类记忆强有力、通用而快速的扩展。

然而，便捷的提取还存在另一个不那么明显、但或许更为麻烦的问题。随着新的数字提取工具克服了传统模拟存储系统连续存储的本性，搜索时间大大缩短了，但是作为显示在我们电脑屏幕上的搜索结果，它们已经与它们当时的语境都脱离了。相比之下，传统的连续式文件检索提供了相当多的背景信息。它可以帮助一个人理解情况如何随着时间变化，而且可以弥补被数字检索直接获得的信息片段所丢失的大量背景和详细信息。很多被用于数字提取的排序结构（ordering structure）忽视了这种背景信息。

这种“去语境化”并不是数字时代独有的现象。美国加州大学欧文分校的教授西蒙·科尔（Simon Cole）充分详尽地解释了，由执法部门采集的有语境的指纹如何被去语境化，变成抽象的数值特征以便于比较、搜索和提取，导致指纹错误识别概率明显的上升。在其他时候，原始的背景信息不仅被剔除掉，而且被另一个不同的背景所替代。

就像美国加州大学洛杉矶分校的地理学教授迈克尔·加里（Michael Curry）告诉我们的那样，邮政地址系统，包括邮政编码，首次从一个基于某个地方特定背景的系统变成了一个基于地理位置的更为抽象的系统。然后，市场营销公司获取了这些去语境化的地理数据，比如，邮政编码，将它们与人口统计学的信息相结合，从而将它们再次置入语境中。因此，比如如果我们今天听到邮编是90210，它让人立刻想到的不

是加州某个地方的景象，而是富人和名人居住区的景象。

数字工具加速了向抽象排序和分类系统过渡的进程，剥离了原始的背景信息。这种清晰而抽象的排序分类在数字检索中更容易实现和体现出来，即便这样一种方法可能对最终的结果不利。近来，全世界的软件工程师急忙改进我们的数字工具，以求（再次）找回少量的背景信息——比如，通过使用元数据和标签。但是目前，我们将不得不与非常便宜、普遍又易于使用的提取工具作抗争，同时也得与它们产生的去语境化和再语境化的结果中固有的、但又很少被意识到的缺点作抗争。

驱动力4：全球性覆盖

廉价的存储器和易于提取的特征使数字化记忆变成了一个强有力的工具，这个工具扩展了人类能够记住的信息。然而，想要获取存储信息的人们仍然不得不去该信息的所在地。全球性的数字网络消除了地理距离的限制；为了获取一个连接到全球数字网络数据库中的信息，一个人仅仅只需要接入那个网络。这从根本上改变了搜索和提取的经济学，因为亲自去某个资料库获取信息的时间和成本被更为低廉的网络通信费用所替代。

商业数据库供应商最先领会了全球数字网络在信息获取方面的潜力。在20世纪七八十年代，他们将自己的信息数据库连接到全球数字网络中，这为他们开辟了新的市场。这些新的商业用户大部分都将不会到很远的地方去搜索和检索一个数据库，而会发现通过更为便利和实惠的远程访问检索信息的价值。互联网只不过是深化了这一变化。远程访问的沟通成本降低到可以忽略不计的程度，将潜在的客户群从商业用户拓展到所有用户。

随着沟通成本的下降，通过全球网络可访问的信息的丰富性也上升了——从简单的文本到高清的视频。仅仅几十年前还只能穿越大西洋去获取信息的人们，如今在自己舒适的办公室或者家里，以微不足道的成本达到同样的目的，而且非常快速，也不用管所谓的营业时间，每周7天第天24小时都可以获取到数据。

20世纪90年代末，世界上绝大多数人，即便是工业化国家的人们通过拨号连接传输接入互联网的速度，充其量也只有每秒大概7000比特。到了2008年，超过3.3亿人拥有了宽带连接，接收数据的速度超过每秒10万比特。为了获得这个翻了15倍的巨大增速，连接速度必须大概每15个月就翻一番，这样便超过了处理能力和存储能力已经获得的显著增涨（每18个月翻一番）。同时，每月的连接费用则保持相对平稳，这造成了通信成本同样惊人的下降。此外，因为几乎所有的宽带连接都提供了

一个平稳的月租费，所以，他们为用户创造了一个更进一步最大化利用数字网络的经济动机。

有三个驱动因素促进了这一发展。

第一个因素是，互联网的包交换（**packet-switched**）结构。与电话系统不同，它直接连接通信双方，互联网上的信息以小信息包的形式传输，这样可以找到彼此独立的、从发送者到接收者的最快路径。这使得用户能够更好地对可用的网络基础设施进行利用。

第二个因素是，铺设了大量适合宽带连接的光纤。在1999年间，据估计，美国每天铺设了4000英里（1英里约合1609米）的光纤。第一次网络泡沫和电信泡沫的破灭，导致了存储能力的过剩，使得高达95%的长距离光纤最初没有被“点亮”（或者说，未被使用）。虽然这严重地伤害了投资者，但是也使得随后的快速发展成为可能。当投资者看到了Web 2.0服务刺激宽带快速增长，大量的资金又被投入进去以寻求宽带基础设施的进一步提升。

第三个因素是，技术促进了在不铺设新光纤的情况下，改装现有的基础设施以满足需求的增长。光被用于在光纤中传输信息。如果有人能够缩短一个光脉冲表征一个信息位所需的时间，那么，就能够加快现有光纤的传输速度。缩短波长以及通过同一根光纤同时传输多极化的光脉冲，则能够产生另一个数量级的速度提升。现在，现有铜线电缆的传输速度，也已经通过改进DSL（数字用户电路）技术的原始版本得以大大提升。至少从中期来看，在成本不变的情况下，存取速度很可能继续提升。当然，这些可获取全球互联网信息，且让人负担得起的丰富带宽，并不意味着每个人都能获得它。发展中国家以及工业化国家农村地区的人们则继续被遗忘，或者以更低的速度和更高的成本接入互联网——这需要许多不同社会做出重要的政策考虑。

取舍之道

对于全世界数以亿计的人们而言，信息获取已不再取决于地理位置或者支付高昂通信成本的能力，而仅仅取决于信息提供者制定的政策和收取的费用。其影响是深远的：连接到全球互联网的数字信息突然间变得有价值了，而没有被连接到全球互联网的信息则已经与被完全遗忘的信息无异。数字化与广泛而廉价的存储器，易于提取的特性，以及全球性覆盖相结合，使得信息在人类能够负担的情况下，永远保持有效，永远可以获取。

当然，我们存储在我们个人硬盘里的信息不会自动地与全世界相连接，即便我们已经接入了互联网。有人可能会认为，是商业的信息提供商，而非个人，提供了他们的数字网络存储。即便这个观点是正确的

（而且我们将会立即讨论它的优点），我们共享的数字化记忆将仍然会有明显的扩展，而且使遗忘变得更难。

原因就是信息经济学。不像大多数物质产品的生产，比如鞋子或木制椅子，信息产品几乎全部的成本都花在了第一个元件的生产上，而随后复制品制造的成本则相当微不足道。

以一部卖座的电影为例：制作它是最昂贵也最困难的部分，生产拷贝给电影院放映的成本则微不足道。对于书籍、音乐，以及信息数据库而言，同样如此。因此，任何人都有一个强烈的动机想要一个信息数据库，以尽可能多的方式去利用它。增值成本[7]与潜在的额外收益相比，基本是可以忽略的。这推动了信息的收藏者让许多人使用他们的信息资源。

由于同样的原因，商业的信息供应商拥有一个强烈的经济动机，他们将甚至不相关的信息数据库都捆绑在一起销售。这可能听起来不太好理解。举个例子，为什么让消费者能够在访问法庭判例数据库的同时，还能访问包含主要报纸及通讯社内容全文的数据库，是商家明智的策略呢？乍看之下，律师搜索判例与记者研究新闻故事似乎毫无共同之处。但是正如亚诺斯·巴克斯（Yanos Bakos）教授和艾里克·布林约尔弗森（Erik Brynjolfsson）教授所说，这个策略是非常明智的，这比单独出售各个信息库的访问权限要更好。他们的论证直截了当：信息供应商并不确切知道他们的客户想要什么信息，而一个信息库越小越专业化，就越难知道它是否能满足客户的偏好。另一方面，一个信息数据库群能越大地满足更大客户群的偏好，就能吸引一个更大的市场。由于信息生产的经济学，给一个人的信息群增加信息集所造成的高成本，并不会完全抵消更高的潜在收入。这导致那些拥有信息的人将信息授权给商业信息供应商，同时对于这些信息供应商而言，他们会向第三方信息提供商寻求授权，以便扩大他们能够提供的信息群。信息市场中的推动力和吸引力相结合，这促成了几千年来我们人类已经习惯了的遗忘逐渐消亡。

让我们重新再审视一下前面的那个假设，即只有商业信息供应商提供在线的信息，而个人则不会。从技术上来讲，没有什么能够阻止个人在线地提供他们的信息。互联网是一个数字网络，其中的每个参与者既可以是一个发送者也可以是一位接收者。消除了信息寻求者和信息提供者之间、读者和作者之间传统上的差异，自早期互联网以来已经成为一个经常被讨论的特性。起初，它的经济效果是有限的，当时大部分人通过缓慢的拨号连接访问互联网，并且信息存储非常昂贵，但是多年以来这种状况已经发生改变了。全世界越来越多的人通过宽带连接接入互

联网，而宽带连接能够提供更高的信息传送率，而且低廉的连接费用激发了一种利用这个顺畅信息通道的动机，这个动机包括与他人分享信息。为了限制网络的潜在负担，一些宽带提供商禁止用户操作网络服务器，此后对于全世界而言，他们的信息资源只是充当了简单的网页。但是，一些专业的网络托管公司每月收取一定的费用在他们的网络服务器上提供空间。由于2008年的激烈竞争，每月10美元将能够获得50G的网络空间，以及1000G的每月信息传输量，这足够托管2500万页文档，数以万计的数字照片，或者十多个高清视频。克服这些限制的另一个方法是，利用点对点（**P2P**）的音乐和视频文件共享软件，以及针对所有其他文件的共享软件。它们的效果很好，有几千万用户已经发现了这种方法，并且宽带提供商已经更难以从技术上阻止用户了。而且，由于现在点对点信息共享已经成为了一种如此普遍的现象，宽带用户仅仅只是期望他们的提供商不要限制他们的信息共享行为。

依靠网络公司最近的一次高潮，个人也可以很方便地与他人分享信息。这是经常被提供商描述为Web 2.0的服务，它们的目的在于促进对基于相互共享信息的在线社交的创造。

网络相册Flickr网站和视频网站YouTube是这种服务的绝佳示例。任何人都可以上传图片和视频，并且可以免费地将它们存储在网站服务器上，同时全世界都可以访问它们。这样，共享信息就变得异常简单。Flickr和YouTube网站背后的公司——雅虎和Google，都是著名的在线信息提供商。廉价的存储器和宽带访问的全球传播，是他们能够负担得起免费提供这些服务的原因之一。他们还希望最终能利用社交网络，因此，在线社交正在形成。或许最为重要的是，这帮助他们强调了他们作为通往全球信息海洋门户的卓越角色。

全世界数以亿计的用户已经拥抱了这些服务和类似地服务——想想社交网站facebook，Myspace，或者twitter吧。对于他们而言，自2000年以来互联网已经发生了深远的改变。正如我在介绍里所说的那样，在**2000**年以前，互联网仅是访问信息的工具，今天互联网还是共享信息的工具。这意味着，用户从被动的接收者转变为主动的贡献者，这与数字化对生产和消费信息自身之间的关系所造成的变化非常相似。

当然，戈登·贝尔，那个把自己的职业生活记录捕捉在数字存储器上的软件工程师，他着重强调，他自己掌控着能否以及如何让他人获取他个人信息的途径。贝尔本人已经决定不让其他人窥探他的信息宝藏了。他解释说，他的生活博客是为了帮助自己记住，而不是为了让其他人访问他的文件："许多人都把他们的生活放到网络上，但我不支持这种做法。生活博客的创建完全是私人化的，是为了帮助个人自己。"他的观

点可能很好，但是正如他自己承认的那样，这种做法并不能代表目前的互联网用户。

正如我已经提到的那样，美国三分之二的青少年使用互联网去创建信息，并与其他人分享信息。这个巨大且还在不断增长的用户群体已经内化了这种运用互联网信息再创造（bricolage）的文化。他们享受着经济学家称之为网络外部性（network externalities）的巨大好处——每个新用户都加入进来共享信息，这增加了所有现有用户的价值。不过，有人可能会说，即便用户将互联网视为一个全球性的、在个人之间共享信息的工具，用户仍然可以选择谨慎地允许他人通过全球网络访问信息。如果他们不想或者不再想让他人访问，他们也可以这么做。毕竟，数字信息之美正在于，清除信息就像点击“删除”键一样简单。这乍听起来好像是对的，然而，至少有两个原因说明，这几乎不可能是对的，这两个原因都是信息网络如何运转的基础。

第一个原因是，一旦一个人已经分享了信息，这个人就基本上失去了对该信息的控制。如果我让其他人访问我的信息，我必须信任他们不会违背我的意愿去使用我的信息。但如果他们辜负了我的信任，我能做的却微乎其微。谁也不想让一些技术高超的黑客闯入自己的电脑窃取信息，然而这却会发生在许多毫无戒心的用户身上，而且它也并不总是需要高超的技术。就像帕丽斯·希尔顿[8]在2005年经历的那样。

通过手机供应商提供的一项功能，希尔顿将手机里存储的信息备份到供应商的中央电脑上。访问存储的信息需要一个密码，或者对于那些忘记他们密码的用户，需要输入他们最喜欢的宠物的名字。

作为一位名人，希尔顿宠物狗的名字是众所周知的。一个17岁的少年运用这个信息进入了希尔顿备份的手机通讯录和数字照片文件，然后将这些信息通过点对点文件共享给了全世界。这个少年随后被捕，并且因参与另一起案件而被审判。但是伤害已经发生了。这不仅仅是帕丽斯·希尔顿的私人信息被曝光了；她手机通讯录中所有人的手机号码，往往还有未编入册的，也都暴露了。许多人不得不换掉他们的手机号码，他们需要将新号码逐一告诉给自己的朋友、同事和商业伙伴。所有这些人都失去了对他们手机号码的控制，不是在那个少年进入希尔顿文件的时候，而是更早，当他们把手机号码告诉给希尔顿的时候。

这可能听起来像一个极端案例，而且手机供应商迅速修补了这个缺陷。但是，如BBC所报道，facebook上有超过1亿毫无戒心的用户可能都存在私人信息被曝光的风险。不妨想想，其实每天都上演几百万次的信息滥用情况：我们告诉某个朋友或同事一些有价值的信息，也许是一个

好主意，一个事实，或者仅仅只是闲谈八卦，他再将这个信息传递给第三个人。他并没有征求我们的同意，事实上，我们甚至都不知道。如果我们知道的话，也只是偶然地发现我们的话被传出去了。在模拟时代，这个危险多少还被个人传播信息的有限能力缓和了。我们很少有人能够控制一台印刷机或者管理一个电视节目。然而，即便这个信息传播地很慢或者仅仅只传给了少数几个人，它仍然能造成相当大的损害。这也是为什么我们不把所有事情都告诉所有人的原因，而且谨慎地决定何时告诉谁哪些内容。

此外，由于在模拟时代大多数这种个人沟通都是口头的，在某些情况下一个人还可以假装那都是一个巨大的“误会”，最终这个事件将会被遗忘。在数字时代，情况更加复杂。信息的全球性覆盖，让它能够更快地在比以前更广阔的范围内传播。数字化使复制和共享变得更容易更快捷，因此也更难以控制，同时不断增加的媒体丰富性（音频、视频以及文本）使信息变得更加难以否认。而且，由于数字拷贝是原始文件的完美复制品，跟踪信息流向某个可能的漏洞，或者识别出它的源头非常困难。在数字时代，召回信息以及阻止他人分享信息，已经变得非常困难，尤其是一旦某个信息已经开始传出狭窄的地理学和人口学范围的时候。

第二个原因是，每次在线交流本身（即便不分享文件），就等于让交流伙伴了解了关于自己的信息，交流伙伴也可能把这个信息拿来与他人分享。这会让我们自己不再是我们信息资源的主人。

如果我在亚马逊网站上订购了一本书，我就留下了一个信息痕迹，亚马逊就可以运用它与其他信息一起，为我推荐书籍。也许有人能够预料到这一点。但是，当我仅仅只是浏览亚马逊的在线商店的时候，同样也是如此，即便我并没有明确告诉亚马逊去监视我的浏览习惯。无论如何他们都会这么做。有时，我会从中受益：当亚马逊根据我已经留下的信息痕迹所推荐的一系列产品吸引我的时候。大多数在线零售商都会这么做，就像大多数公司都会提供基于网络的电子邮箱、社交网络和信息共享一样。同时，互联网搜索公司会仔细地监视用户在搜索什么内容。

有些人可能反对我们并不必去实体店或者可以在线购物的看法，反对我们发送或接收电子邮件、写博客、在虚拟世界里工作、使用搜索引擎，或者允许cookies（信息记录程序）。他们认为，这仍然需要我们的参与，其他人才能记录我们的数字踪迹。

但是，这个论证明显是不诚实的，至少有两个理由。首先，往往用户并不知道他们的数字活动正在被记录并被提交到数字存储器中，从浏览亚马逊网站到使用Gmail起草电子邮件。类似地，当输入一个网址

或者填写一份订货单的时候，用户可能很享受他们的网络浏览器中的这些自动完成的功能，但是并不清楚他们是否理解了，这也意味着他们的网络浏览器会保存他们的活动记录。用户可能喜欢他们现在能够运用强大的全文索引和检索快速搜索他们硬盘的感觉，但是他们可能没有理解，这是因为他们存储在他们个人电脑里的任何内容都被默认地编入了索引。断开这些记忆的常态很复杂，而且在许多情况下也是不可能的。其次，即便用户清楚地知道他们的活动被持续而广泛地记录为数字化记忆，他们唯一可用的简单补救方法，将是把自己变成“数字隐士”，戒掉在线聊天，同时还要戒掉许多离线的交流。这是因为，每次与世界的交流都会产生一些信息。通信理论家保罗·瓦茨拉维克（Paul Watzlawick）曾经在一个更为普遍的背景下指出，“一个人无法不交流。”而且这种情况下，只要是在沟通过程中运用任何数字设备，它将有可能被添加到我们的“数字人生”中。

当信息被添加到这种由多种信息源混合而成的数字网络中，信息的重新组合可以创造任何东西，从完全平庸的到具有难以置信创造力的，从有趣的到深刻的，从简单的到复杂的。我们的个人电脑逐渐充当了多功能信息处理器，我们的硬盘可以提供充足的存储空间，精细的编辑和混合软件能够让更多的人负担得起，人们可以很容易地使用、剪辑、编辑和添加。而且，如今对于成千上万的用户而言，将数字信息重新组合都变得触手可及了。但是这些新文化产品的创造者将信息与其已有的背景剥离开，并将它与一个全新的背景相联系。俄罗斯的“妖精编辑”，之前我们提到过，他们给卖座的电影重新配音，它之所以如此非同寻常，恰恰是因为他们用一段全新的对话替代我们熟悉的台词，将我们以为我们知道的电影重新置入新的背景中，并且给了它们一个非常不同的电影镜头。音乐采样，即重新利用著名音乐的片段作为一张新唱片的元素，也是一种类似地概念。

美国堪萨斯州立大学教授迈克尔·韦施（Michael Wesch）的视频《*Web 2.0*》提供了一个非常吸引人的例子。同时，这也是数字化再创造文化兴起的一个完美而具有说服力的解释。

2007年年初，他将视频上传到YouTube网站上。几个星期之内，已经有几百万人看了这个视频。这在模拟时代是不可能发生的事情。但是更为重要的是，在随后的几个星期中，有人将韦施的视频翻译成了不同的语言，有人制作了评论视频讨论韦施的一些设想，但那些评论视频也在模仿他的艺术风格，甚至完全对韦施的原始作品进行了重新混音。韦施的视频变得很有名，不仅仅是因为有几百万人看了它，更是因为许多人将它作为他们自己作品的起始点。

对于那些跟踪YouTube以及众多其他信息共享网站上发生了什么的人而言，韦施的视频以及它所触发的故事他们非常熟悉。通过重新组合他人已有的信息而不仅仅是被动使用自己事先记录的内容进行积极创作的能力，已经触发了巨大的创造潜能。自然而然，这也造成了被共享的数字信息在原创作者无法掌控的情况下，出现被去语境化和再语境化的局面。

取舍之道

信息提取已经造成了信息的去语境化和再语境化，因为信息片段是在剥离了它们相应背景的情况下被提取的，而且被呈现在众多搜索结果构成的新背景中。不过，花上足够多的时间和努力，一个人往往能够从搜索到的结果追溯到信息的源头，从而在最初的语境体验它们。这就像使用索引在一本书中寻找某个段落——一个人仅仅通过阅读一个句子或者一页内容可能无法掌握完整的背景。但是，如果内容足够重要，那么他们就可以阅读整个章节甚至整本书。通过数字化再创造，这变得不再可能，因为组成再创作的作品的信息片段被从它们的原始语境中抽离出来了，没有办法追溯它们的信息源。

假设一张照片被用在一张幻灯片中，一个句子摘录自某个人的小说，或者某些调子借用了某首歌，在这些数字再创作的作品中我们没有可参考的信息去追溯“源头”，因此那张照片最先被拍摄时的背景，那个句子被写出来时的背景，或者那个音符最先被创作和演奏时的背景都无法追溯。如果什么也追溯不到了，就加大了误解的危险。

因此戈登·贝尔可能不会泄露他已经在他的数字笔记中存储了什么，但是如果他参与互联网中信息创造与共享的文化，他将会发现随着时间的推移，他数字笔记中的许多内容可能出现在其他人手中。有时，他可能会谨慎地分享信息，信任他的信息接收者。但在平时，他可能没意识到或者已经忘记了他的在线交流所产生的信息，已被其他人所存储和运用。请不要误解我：共享信息常常是极为有益的，即便在模拟时代，完美的信息控制也只是一种幻想，只有那些最坚定的隐士才能实现。然而，在数字时代，作为个人，我们控制信息的能力被再次大大减弱了。

毫无疑问，过去60年间的技术突破，促进了对数量不断增长的信息的存储和提取。这大大降低了加工、存储和提取数字信息的成本，也成为了很多人共同使用新工具强烈而持久的动机。或许同样重要的是，数字存储器如何为我们提供了一种延续和保存的策略，以超越我们个人必然会死去的命运。正如哲学家罗伯特·诺奇克（Robert Nozick）所说，它让我们留下痕迹，从而意味着“一个人的生命拥有某种意义”。因

为使用了数字存储器，就算有一天我们去世了，我们的思想、情感和经验依然不会遗失，仍然能被子孙后代所运用。借助这些，我们能够继续活着，逃脱了被遗忘的命运。随着现代社会生育率的垂直下降，除了我们想要与子孙后代共享信息的传统背景之外，保存我们生活记忆的渴望可能会更加强烈。这是一种非常人类化的策略，以确保我们没有白活一场，以及我们在死后不会被迅速遗忘，就仿佛我们从没来过这个世上一样。

广泛的数字化，让越来越多的信息被转化为一种常见的二进制代码。通过廉价的存储技术，保存数字信息已经不仅变得可以负担得起，而且往往比选择性地删除一些信息所耗费的时间成本更低。简便的提取工具，也使得我们能够史无前例地利用我们巨大的数字信息资源。同时，通过快速又负担得起的数字网络实现了全球性覆盖，这已经不仅使得远程访问这些信息资源成为可能，而且促进了创造、再创造和共享的文化。

在这种文化中，我们已经抛弃了传统的信息控制形式。同时，随着更多的人在寻找除了生育之外超越我们必死命运的策略，我们人类对更加完整数字化记忆的需求将会继续上升。这将导致整个世界被设置为记忆模式，而且即便有的话，也只有很小的动机去遗忘。

---

[1] 赖特为20世纪上半叶最有影响的建筑师之一，影响了整个美国建筑的发展。——译者注

[2] 该文标志着情报学的诞生。——作者注

[3] 艾字节（exabyte）相当于10亿吉字节（gigabyte，GB），或者100万太字节（terabyte，TB），等于10亿个字节的10亿倍。——作者注

[4] 校验和（checksum），指的是在数据处理和数据通信领域中，用于校验目的的一组数据项的和。这些数据项可以是数字或在计算检验的过程中看作数字的其他字符串。——编者注

[5] bricolage，指使用手头现成工具制成的物品。——作者注

[6] metadata，与目标文件相关的一个数字文件的数据。——译者注

[7] 增值成本一般是产品生产中必需的成本支出，如产品加工过程中所消耗的材料、人工、能源、设备损耗等。——作者注

[8] Paris Hilton，美国希尔顿集团继承人之一，女商人、模特、时尚设计师、歌手、演员及作家。——译者注

04 一个没有安全与时间的未来：数字化记忆的两大威胁

在信息权力与时间的交会处，永久的记忆创造了空间和时间圆形监狱的幽灵。完整的数字化记忆摧毁了历史，损害了我们的判断和我们及时行动的能力，让我们无助地徘徊在两个同样让人不安的选择之间：是选择永久的过去，还是被忽视的现在……

从某种意义上看，遗忘并不是令人困扰的缺陷，而是一种足以救命的优势。当我们忘记了过去的时候，我们便重新获得了去概括、去概念化的自由，以及最重要的行动自由。

数字化记忆加深了信息富民与信息贫民之间已经存在的鸿沟，进一步增强了权力的倾斜。

没有了某种形式的遗忘，原谅则成为一件非常困难的事情。如果所有的时间都永远是现在，那么所有的时间都不能够得到拯救。

遗忘终止的未来

人类渴望记忆，不过在通常情况下他们会遗忘。为了减小这种生物学上的局限，我们开发出了很多工具，从书籍到视频，借此充当我们的外部存储器。事实证明，这些工具非常有用，因为自从有了它们之后，人们比以往任何时候都更加容易记住。就在几十年前，这些工具还没有打破记忆和遗忘之间的平衡：记住只是例外，而遗忘则是常态。

在数字时代，这种平衡已经从根本上被改变了。数字化——数字革命的理论基础，带来了廉价的存储器、便捷的提取，以及全球性的覆盖。如今，遗忘已经变得昂贵而又困难，记忆反而便宜又容易。在数字化工具的帮助下，无论是从个人还是社会层面，我们都已经开始抛弃遗忘，开始从我们的日常生活中抹去一种人类最基本的行为机制。

从某种程度上讲，这是值得庆祝的。毕竟，人类已经努力了几千年，来克服生物性遗忘的约束和人类记忆的局限。从记忆准确率和效率的提高，到超越必死命运的希望，海量且容易获得的数字化记忆为我们提供了数不清的好处。

举个例子，包含了某个人多年前罕见过敏史信息的全面数字化医疗档案，也许能挽救一个人的生命。与其他方式相比时，它显得更加有效，比如说依靠某位医生不完美的记忆，或者多年后已经无法看懂的手写医嘱进行治疗。同样，相比于目击者大脑对事发当时的再现，闭路电视记录和飞机黑匣子能够为事故调查者提供更加准确的信息。

全面而便于查阅的记忆带来的额外效率同样是非常可观的。更快和更全面的信息传播（无论是科学结果还是原始数据）都有可能促进创新

和经济增长。更精确的数字化记忆也许能够帮助企业更精准地设计和生产出人们喜爱的产品，并把产品推向更精确的目标人群，从而将企业、消费者和市场环境从几百万吨令人讨厌的营销和广告材料中拯救出来。消费者也不喜欢以这样的方式成为目标受众。

经济学家一直认为，更好的信息可以提高效率。eBay的买家和卖家评价系统就是一个好例子。它作为所有市场参与者全面的外部记忆，为未来潜在的交易伙伴提供了关于对方历史行为有价值的信息。但数字化记忆所带来的好处并不只限于提高了市场交易的效率。它们还延伸到更广阔的社会领域，以及那些过去有记录但又不易获取，从而无法被他人所利用的领域。

纽约市健康与心理卫生部（New York City Department of Health and Mental Hygiene）曾对纽约市的餐厅做过一次卫生审查。一旦这些审查信息可以上网查询，那么对于有食品卫生问题的餐厅而言，这会立即影响它的声誉。

类似地，多年来美国环境保护署（EPA）都要求污染者自己报告毒素与污染物的存储量和排放量。这些信息与相关的信息一起，都原原本本地录入进了一个数字化数据库，但是仅此而已。只有当环保署使数据库可以网上查询，并将它和易于使用的在线地图结合在一起，这样消费者（而非研究者）才能轻松地查询他们附近的污染源，促使他们参加当地的清洁活动。

此外，环保署像纽约市健康与心理卫生部一样，都应该为包括非政府组织在内的第三方机构开放数据库，并且将第三方机构反馈回的信息与他们自己原来的信息、地图和人口学资料整合起来。

正如玛丽·格雷汉姆（Mary Graham）所言，这种公开示众的确能够对行业行为产生影响。不管是eBay上的交易信用，还是纽约市健康与心理卫生部或者美国环保署的环境信誉，没有了这种全面的数字化记忆，它们都无法正常运作。

数字化记忆可能还能提高政府在个人和社会层面决策的准确性。当个人申请某项政府服务时，全面的数字化记忆也许可以让政府机关做出更准确的决策。在社会层面上，这些易于获得的数字化记忆有助于预测一般趋势和社会发展，使得政策制定者在问题失控前可以调整政策。此外，选民可以访问数字化记忆，看看他们的政治代表是否以及如何恪守他们的承诺。

从遗忘转换到完美的记忆，这些都是我们应当感到欣慰而非担忧的常见理由。不过，哪怕只是稍微详细地审视一下完整的数字化记忆所带来的好处，就会发现事情没那么简单。比如说，eBay上的信誉并不像信

息经济学家所认为的那样，准确地反映了交易的满意度。研究者发现，eBay上的卖家会策略性地提前他们评价买家的时间。许多卖家甚至在交易结束之前，就给买家很高的评价，而在这时，卖家才刚刚获得可以评价买家的相关信息。他们这么做并不是因为对交易本身很满意，而是因为他们希望买家同样也给他们好评作为回报。

但是，这还只是eBay的客户试图影响eBay交易信誉的数字化记忆的众多方式之一。到2008年春天时，这种普遍的赌博性信用记忆行为促使eBay对它的评价政策进行了大幅度的调整，卖家对买家的评价受到诸多限制。然而，从一开始，eBay这个品牌一直为客户记录历史信誉，它现在已经在慎重地引入对糟糕购物体验的遗忘。

可以看看美国环保署协助整合地理和污染信息的网站。这个网站一直运行良好，直到环保署意识到，其中关于美国有毒废物存储量和排放量全面的数字化记忆，可能会成为恐怖分子寻找合适袭击目标的重要的工具。正如环保署所发现的那样，数字化记忆在一方面带来好处的同时，可能在另一方面会被证明有害。因此，环保署迅速缩小了完全开放访问环保信息数据库的范围。

数字化记忆可能会为我们提供重要的好处，但是正如上面两个例子所展示的那样，它并不一定总是会带来好处。在某些情况下，人们可能会为了他们自己的目的在赌博性欺诈或篡改数字化记忆的行为上取得成功。在其他情况下，易于获得的数字化记忆可能会提高短期的效率，但却会将个人或社会暴露在潜在的恶果面前。

数字化医疗记录的案例指出了另一种可能的困境：假设某个人全部的医疗信息都存储在一个可以访问的数据库里，那么它应该如何组织，才能确保在最可能挽救生命的各种信息已建立起联系时，正在访问它的医生也能够迅速找到相关的各种信息呢？

假设经年累月收集起来的个人信息的背景已经极大地被改变了，那么，被加入到这个人一生当中的信息该如何解释呢？比如，当一个医生面对某种特定疗法或药物“让人困扰的反应”被别的医生描述为“正常反应”，或者甚至在多年后，更有经验的自己认为是“正常反应”时，他会作何感想？

完美的记忆使我们暴露在过滤、选择和解释的挑战面前，而遗忘通常会使我们免于这些挑战。就像eBay和美国环保署一样，我们的医疗服务人员可能也会发现，记住一切的能力并非时时刻刻都是有益的。

在股票市场中，全面的数字化记忆并不能实现所有预想的好处；有时候，更高的准确率和效率带来的好处却被一些经济劣势或伤害抵消了，正如美国环保署的例子。但是，数字化记忆的问题比这更为严重，

而且简单的成本/效益分析并不能轻而易举地缓解这个问题。

取舍之道

几千年来，人类一直都生活在遗忘的世界里。个人的行为、社会的机制与进程，以及人类的价值观都融合在一起，并且反映了这个事实。如果认为在数字化和技术的帮助下，丢弃人类本性中这个根本的部分并不是一件痛苦的事情，未免有些太幼稚。人类可以通过多种方式快速适应不同的环境条件，但是人类行为根本特性的改变或替换，则需要几代人的时间。即便我们能够应付一个记忆成为默认状态的新世界，并且熬过痛苦的适应过程，我们就能视其为一个重大的进步吗？或者说，这其实是一个可怕的诅咒吗？

支持记忆的论据都集中在这些观念上：以数字形态被记住的信息具有能感知得到的高质量，运用数字化记忆所产生的高效率，以及数字化记忆承诺提供的更卓越的信息不朽。这些论据所忽视的，却是用记忆替代遗忘对人类与整个社会的真正影响。我们的生活将会受到怎样的影响呢？人类的行为将如何调整，结果是能够超越我们目前现状的提升，还是换来更多的损失呢？

一个没有遗忘的世界是很难预知的，但是基于已有的研究、严密的分析和一些推测，我们还是能够用两个术语勾勒出当前所面对威胁的轮廓：信息权力和时间。

信息富民VS信息贫民：信息控制权的威胁

“知识就是力量。”四百年前，培根曾这样写道。这是一个早已深深植根于人类行为中的观念：获取关于是什么围绕在我们身边的信息，能够帮助我们生存下来。在很多情况下，这种获取信息的能力都是相对的，源于我们比他人知道得更多。从人类早期开始，应运而生的特殊社会等级、特殊群体，和特殊职业的人总能从特殊的渠道获取信息。这些人可以与神对话，治愈疾病，控制精神，指挥军队，改进产品，统治社会。而当我们那么好奇，那么想知道我们身边发生了什么，我们就会不自觉地不去与每个人分享全部的信息。

将信息保护在我们自己的范围内以保留信息优势的行为可能是通过文化习得的，而且也许是近几十年来发展的结果。不过，这一行为很可能有更古老的根源。

16世纪席卷欧洲的宗教改革运动削弱了教堂作为社会核心信息媒介的作用，并由此部分打破了天主教会的势力。在马丁·路德之后，人们可以自己阅读《圣经》，而不是被动地接受某些“知情人士”告诉他们的内容。而且，相应地，人们也可以直接向上帝而不是教堂的神职人员忏

悔自己的罪恶，因此，剥夺了教会在控制重要的内幕信息方面的权力。结果是天主教会的势力被削弱了，非宗教的民族国家获得了崛起的机会。

以书籍为例，当印刷机广泛传播后，许多国家的图书出版商与政府联手，建立起信息控制系统以巩固他们自己的权力。只有通过政府的审查，书籍才能得以出版。作为回报，出版商工会获得了政府的授权，拥有书籍的独家印刷特权。就在此刻，安妮女王（Queen Anne）认为图书出版商的权力过大，第一部授予作者对他们著作所有权的立法（1709年的版权法案）可以说就出台了。

“版权”（Copyright），这个由安妮女王发明的法律术语，有望通过赋予作家权力以打破出版商的强势。正如罗伯特·基欧汉教授（Robert Keohane）和约瑟夫·奈教授（Joseph Nye）所解释的那样，从很早之前人们就已经理解，信息的力量源自于对信息获取的控制能力。当然，并非所有人都想当作家都会斟酌是否要出版自己的作品。但是，所有人都拥有关于自己的信息：我是谁，从哪里来，哪些信息可能想与他人分享或者不想与人分享。问题的关键就在于，人们能在何种程度上控制他人获取关于自己的何种信息。

长久以来，“保持距离”一直是对付吵闹的邻居最有效的办法。如果与自己最近的农场也在好几英里之外，那么一个人做什么或说什么都很难被传播开，除非有意为之。而另一方面，在城市里，人们很难避免被他人观察的命运。但同时，城市的大量人口也保证了一定程度的匿名性。当每个人被信息轰炸时，大部分信息被淹没在城市生活制造出的，感官印象构成的海洋里。在这两种情况下，个人都保持住了，至少某种程度上保持住了，他人可以获取关于他们的何种信息的控制力。相比之下，在乡村以及类似地社区中，居民数量太少、彼此之间的距离太近，使得匿名的面纱无法发挥作用，控制一个人的私人信息流动也变得更为困难。在这种情况下，对个人信息的控制可能只能来自于社会规范——比如，私自阅读他人的日记通常被认为是不合适的。

如果其他所有控制个人信息的方法都失效了，人们依然还有另外一种选择，尽管这个选择代价更为高昂：那就是离开。几个世纪以来，从一个社区搬到另一个社区允许人们以清白的历史重新开始他们的生活，因为关于他们从前的信息仍然留在当地。在18世纪和19世纪，横渡大西洋从欧洲来到美国新大陆，使得人们可以从头开始，不仅是经济上，更重要的是在他人关于自己的评价上重新开始。甚至从一个大城市的一个社区搬到另一个社区也可能产生相似的效果。从某种程度上说，这种搬家无异于宣告信息破产；但在此之后，人们可以在控制自己的私人

信息方面重新再来一次。不过，这种重新开始的成本，包括通过地理搬迁实现的重新开始的成本是巨大的。此外，成本还取决于某人与他所在社区的联系。一般而言，那些与原社区有很多联系的人要比那些与原社区几乎没有联系的人面临高得多的代价，也比那些只与几个社区有联系的人代价要高许多。所以，最后面那些人想要离开，只需部分地切断而无需完全切断他们与外部世界的联系。

在过去的几十年中，我们已经目睹了人们对其私人信息逐渐失去控制的过程。企业拥有明显的经济动机去了解更多关于其现有客户和潜在客户的喜好信息，以及确保他们之间的交易高效进行。福利国家拥有精准地找到公共服务目标人群的渴望，而我们的社会也渴望降低公民的人身风险，这使得更多的私人信息必须从个人向社会和私营部门组织流动。数字化与廉价的信息处理和存储相结合，已经使得信息的流动在技术上和成本上都成为可能。这让一些担心这种发展会威胁我们个人自由的评论家感到很不安。

20世纪60年代，当第一代数字存储器风靡美国的时候，诸如艾伦·威斯汀（Alan Westin）和亚瑟·米勒（Arthur Miller）等学者坚决反对，并称其为“对隐私的侵犯”。他们所使用的术语——信息隐私（information privacy），一直沿用至今。尽管个人想要控制其私人信息的的渴望已经远远超出了我们通常所说的“隐私”这个词的狭窄定义。

在数字化趋势发展了40年之后，一些人认为我们已经适应了新的数字时代，并已理解了当我们与他人分享我们的私人信息时，我们日常的许多交往中所隐藏的风险。但我对此持怀疑态度。因为，我们仅仅善于判断对我们来说显而易见的事情。不幸的是，记忆的默认状态，对于我们保持对信息控制的潜在影响（而非因此掌握权力），很少是显而易见的。起初我们还没注意到多少，这些影响已经发生了。我们很可能在意识到之前，就已经遭遇到了对信息控制的减弱。由于同样的原因，其他人从我们丢失的信息中获得了信息权力，这会影响我们未来与世界沟通的情形，以及我们如何作为社会中的一员。数字化记忆的三个特征——可访问性、持久性、全面性，使得这种影响成为可能。

可访问性

让我们认为自己能控制私人信息的一种方法是，可以根据接收者与情形的不同来决定是否分享它。比如，在手术前我们会告诉医生自己患有乙肝，但是我们大概不会把这个信息告诉理发师。而且，如果只是在网球场上偶尔遇见自己的医生，我们可能也不会告诉他。我们还期望我们的医生能够严格按照我们与他分享这个信息的初衷去使用它，那就是采取一种合适的医疗手段来治愈我们的疾病。再抽象一点说，我们是基

于特定的语境和目的才愿意与他人分享信息的。为此，我们已经权衡过这种信息量相对损失所带来的利与弊了。如果信息是被单独地存储在互不相通的“容器”中，无论是旧式的档案柜还是自己医生电脑上的数字化数据库（医生的个人电脑没有联网），这都会进行得很好。照此来看，即使在数字时代，我们仍然可以通过只将一小部分信息分享给不同的接收者，以保持对我们私人信息微弱的控制力。然而，这种实际上作用并不大的策略，被人们今天收集、访问和传播信息的广泛能力渐渐破坏了。而正是廉价的存储器、便捷的提取和全球性覆盖才使得这种能力成为可能，这三个特点正是我在第3章中描述的完整数字化记忆的核心元素。

当一个人给美国邮局提交一份地址变更表，他会认为邮局将只在内部使用这个信息，以更改他的邮箱地址。但是事实却并非如此。邮局还会让第三方机构（主要是邮件目录公司）访问地址变更信息，所以，他们便可以更新他们的直邮广告数据库，以保证无论你搬到哪里，最新的户外用品品牌里昂·比恩或是运动品牌兰茨·恩德（Lands End）的商品目录都能寄给你。即便有的话，4000万每年都会变更地址的美国人中也只有少数人知道，他们新的地址信息一直被第三方机构以这样的方式访问。

更令人不安的是，在美国，实际上有三分之二的健康保险公司会通过访问客户的数字化处方历史来筛选健康保险申请人。大多数申请人甚至是许多保险经纪人，都不知道这种通过数字化可访性实现的侵权行径。

当史黛西·施奈德把那张名叫“喝醉的海盗”的照片放到她的MySpace上时，她真的预料到自己今后的主管日后会介意吗？不妨看看照片分享网站Flickr。它不仅允许用户上传照片，而且还可以给照片添加标签。这些标签使得在Flickr的20亿张照片中搜索成为可能。你可能没有上传过任何有损你名声的照片，但是其他人可能上传过。有了Flickr网站上的照片标签，找到一张有你的照片变得容易了许多——而且照片是面向整个世界的。在模拟时代，这一切还都不可能。你可能已经把自己多年前某张有伤风化的照片给别人看过，但是除非你是一个公众人物，否则照片不会被广泛传播。有了数字化记忆后，在人们对于信息仅有的控制力中一个重要部分消失了——成千上万能够上网的人都可以查到你曾上传过的信息。

视频监控的发展为我们提供了第三个恰当的例子。英国安装了几百万的闭路电视摄像头来防止犯罪。但是正如研究所显示的那样，犯罪率只下降了一点点；人们发现，警方并没有足够的人力去浏览这些监控器

拍下的录像。然而被保存起来的话，这些信息又变得不可访问。伦敦警方现在正在尝试用升级后的软件去识别衣服上的不同商标，以求认出其穿戴者，而放弃采用低准确率的面部识别技术。因此，这可能为法律的执行提供了史无前例的可访问性，访问锁定在他们视频图书馆中有价值的信息。

持久性

数字化记忆的第二个特征就是持久性，这个特性能够导致信息权力的再分配。在数字化记忆时代之前，大部分的信息，甚至是犯罪信息，都相对快速地被遗忘了。在一些特殊的案例中，为了保存记忆而付出的努力是必要而且昂贵的。比如在苏联，克格勃（KGB）会在政治犯的档案上盖上“永久保存”的印迹。这表明，永远不能忘记反对者的身份、信念、行为和言语，尽管要做到这些需要付出很大努力。在数字时代，对记忆的努力在逐渐减少，但价值却在提升。Google记忆住所有的检索需求，尽管这些检索需求每天都会出现。如果有人看着你在网上浏览，几天后这个人就会忘了你曾检索些什么，什么时候检索的。但是Google不会，它对个体检索需求的记忆更持久，更精确。

可访问性和持久性结合在一起，使得人类再也不能成功地逃离他们的过去。过往信息将紧随他们，随时都可能被他人通过互联网挖掘出来。

全面性

第三个特征是全面性。通过可访问性和持久性，其他人可以到达存储在日积月累的各个不同数据库中的个人信息。然而，信息提供者可以，而且已经开始更进一步来整合这些来源各异的个人信息，以便在几秒之内建立有关他们、他们的活动、朋友、偏好以及信仰信息丰富的档案。

商业实体已经为几百万美国人提供了包含超过一千个个体数据点/人的卷宗。那些不愿意花很多钱的人，也可以查询他们的邻居、朋友、约会对象或商业伙伴。数字化记忆的经济使得建立个人卷宗更便宜。特别是在连接社区或社会网络的希望驱动之下，我们中许多人已经在线添加我们的个人简介。索引个体的搜索引擎已经出现了。他们使用免费的公共资源，实时塑造了迷你卷宗。其中一个名为Spoct.com的搜索引擎自称拥有“索引超过1亿人，包括超过15亿条的数据档案”。这些针对个人的搜索引擎中，大部分依然十分粗糙且不完整，然而这一情况很快得到改善。同时，传统网络搜索引擎不断将他们的索引能力从传统网页扩大到社交网络，以及许多其他的公共可用数据库。一个人依次得到所有这些信息源，并编辑一个完整的卷宗需要很多时间，然而整合的搜索引擎

使得这一活动变得更快、瑕疵更少、更简单、花费更少。

值得注意的是，从这些古怪的搜索中浮现出来的（即使其中的所有事实都是完全准确的），不是对一个人当前本质适合的解释，而是关于我们生活的一种古怪的人工合成物，它只包含以数字形式可获取到的信息，省略了其他所有信息。

还记得安德鲁·费尔德玛的案例吗？这位加拿大心理咨询师被拒绝入境美国，因为一位边防警察使用Google在网上发现他曾在40多年前服用过致幻剂。直到最近，人们才可以合理地期待多年前出于不同目的给予不同接收者的信息，将能在很大程度上对第三方机构保持不可访问性。当然，我们的医生可能会和我的理发师一起打网球，并且偶然告诉他我的病史。尽管这种偶然的泄露事件会发生，但是在大多数情况下它们的影响是有限的。然而，在数字化记忆时代，这些泄露已经变成了常态而非例外——而且它的后果无处不在。

全面性同样会有错误信息推断的危险。当你在亚马逊上浏览和购物时，就会有很多书或者其他的商品推荐呈现在你面前。亚马逊没有读心术，但是他们会利用你浏览和购买物品的倾向来搜索在他们资料中其他有类似喜好的买家。如果他们定位到了一个人，他们就会把那个买家买过的其他产品推荐给你，因为他们假定你们买的共同的产品能反映出你们有共同的喜好。问题在于，这种信息推断是以概率为基础的。他们精细地把个人和特定的喜好联结起来。当这种基于个人资料匹配的推断用于实践时，结果就是一个扩大的数字版过失。

大多人都不喜欢毫无意识或意料之外地对自己的信息失去控制。如果在和他人面对面沟通时，有人能够强大地运用他们会讨价还价的优势地位，在不必分享他们自己信息的情况下获取他人的信息，将会发生情况呢？为了阻止这一危险的发生，小说家大卫·布林（David Brin）提出了“互惠的透明性”（reciprocal transparency）原则，也就是只有在他们愿意互换信息的时候，我们才会与他们分享信息。他希望能用这个原则去克服信息权力中的不平衡，并且帮助我们获得真正的信息对称。

乍一看，这像是一个很好的想法，但是我并不清楚它如何才能防止冲动的生意伙伴与这个系统博弈，就像eBay信誉机制的案例清晰展示的那样。当然，即便它不能防止个体给彼此提供虚假的信息，互惠的透明性也许仍然是对社会有益的。

考虑一下在模拟时代，如果一个人有阴暗面，他还可以把它隐藏起来。如果你不能很好地隐藏它，那么你就要小心了。法律历史学家劳伦斯·弗里德曼（Lawrence Friedman）在他关于名誉的权威著作中这样告诉我们。如果人们不再能够轻易地隐藏他们的黑暗面，如果我们更加或

者完全了解彼此，那么这个原则不就是对社会有益的吗？如果平等地获取信息取代了我们对个人信息的严格控制，那么欺诈难道不会变得更难吗？布林的愿景是这样一个世界，在这个世界中，分散、互惠地获取个人信息会逐渐攻破那些可以迫使（而且已经迫使了）他们获取社会信息的束缚。

不幸的是，人类记忆并不是一个从过去中机械提取事实的过程。丹尼尔·沙克特（Daniel Schacter）曾对此进行过充分论述，人类记忆基于现在对我们过去不断的重构。正如他所说，“在确定什么是记住的什么是过去真正发生的这个问题上，现在的影响起着更大的作用。”

取舍之道

当我们不断忘记和重构过去的元素时，他人利用数字化记忆可以获取没有被重构过的事实。因此，我们所记得的过去一直在缓慢地改变和发展，而在数字化记忆中捕获的过去却是恒定不变的，定格在过去的时间上。很可能这两种愿景——别人关于我们的凝固的记忆与留在我们脑海里不断发展的记忆，将会发生冲突。两者都不是一个精确完整的关于我们是谁的描述：前者是定格在过去时间上的，后者是我们的大脑对过去的演绎，会被当前我们是谁所深深地影响。

在这样的世界中，我们是仍然求助于人类的记忆，还是把目光投向数字化记忆以确保他人不知道更多关于我们的信息？就算是回忆我们非常私人的过去时，我们是求助于人类自己的记忆，还是转向数字化记忆呢？如果我们所有人都从人类个体本身转向一些外部的数字化记忆，并就此放弃通过我们现在的眼睛去看我们过去的能力？这难道不是与布林极力保留个体控制信息能力的想法背道而驰吗？

布林这个令人着迷的观点还有另一个缺陷。他假想了一个信息对称的世界，个人拒绝单方面转让对他们信息的控制权。但是，这个观点暗示了，“互惠的透明度”的转让，可以在平等的基础上进行，而且个人不会被强迫自己支付给强大生意伙伴的比自己得到的更多。这个想法太理想主义了。

事实绝不会是这样一幅美好的图景——个人通过公正平等地转让对他们信息的控制与访问权而获得信息对称。在现实中，强大的生意伙伴（比如大型公司或者政府）可能会利用信息权力差异来获得信息优势。我们可能看到的不是信息互惠，而是信息权力接踵而至地从无权者流向有权者。如果把这些放到信息隐私的背景下来讲，就是从被监视者转向监视者。然而，这些通常是在未经无权者明确同意（或者无权者未被告知）时就已经发生了。这种信息再分配本身也令人深感不安（第5章将会详细叙述），因为它放大并加深了信息富者和信息贫民之间现已

存在的信息权力差异。

对于我们如何看待自己过去的长远影响，同样令人不安：因为我们不停地在记忆中构建和重构我们的过去，久而久之，我们对这种过去与现在交互作用的回忆，将在质量上变得逊色于拥有我们丰富信息的数字存储图像，从而基本上否认了我们自己的过去。

取舍之道

信息隐私专家长期以来一直认为：对自己的信息保留一定的控制，可以给一个人留有一些他需要的空间来定义他自己。正如我之前所说的，如果数字化记忆的广泛应用导致了信息控制的减弱，它损害的恰恰是这些专家一直在讨论的塑造自己身份的自由。除非一个人是在数字化记忆的世界中丧失了对信息的控制，否则他相对信息权力的损失可能比传统意义上信息隐私概念所包含的内容要更大一些。

目前为止，我们已经考察了在二元关系中数字化记忆如何改变信息权力的平衡，以及它是在何种条件下怎样导致了交易的发生。比如，消费者可能会因为信息灵通的卖家利用数字化记忆所持有的信息差异，而愿意为商品和服务支付更多钱。

想想全球汽车事故数据库；那些租车的人会根据他们过去的驾驶历史而被收取不同的费用，而借助数字化记忆可以立即从租赁公司那里获取他们的驾驶历史。如果驾驶历史确切地表明他们过去曾经出过车祸，那么他们也仍然不得不支付更多的租金。避免支付额外费用的唯一方法是数字化记忆时代可能不再有的东西：那就是遗忘，或者至少从单方面“重构”一个人的个人历史。但是，如果我知道事实发生后我无法修改我的驾驶记录，那么为了避免事故继续影响我的驾驶历史，我将会更为小心地开车。如果过去是不可能赎回的，那么我们唯一的选择就是对未来更加谨慎。

仅对一次特殊的交易而言，信息权力的改变没有产生后果，而不管后果是否以及在何种条件下发生，它都有可能影响人类的行为。还记得那位满腔抱负的实习教师史黛西·斯奈德和心理咨询师安德鲁·费尔德玛吗？事后他们两人都说，如果重头再来他们会以不同的方式行事。如果他们知道后果会是这样，他们当初就会对自己进行自我审查。但是，史黛西与安德鲁的之间存在一个重要的区别。史黛西·斯奈德未能理解数字化记忆可访问性的这一特征——因为某一个目的而向某一位接收者开放的信息，可能会被其他人获取并被用作完全不同的目的。如果了解了这一点，她可能就不会将那张照片上传到她的网页上。重点在这里：总之那些信息可以在史黛西想表达她自己的目的之外，在到达最先的接收者之外，被其他人以不同的理由获取。如果一个人不知道自己的言论

将会以何种方式被何人所利用，那么他就必须做最坏的打算，换句话说，任何批评最终都将会造成极大的伤害。

在与《纽约时报》专栏作家托马斯·弗里德曼（Thomas Friedman）的一次对话中，Google首席执行官埃里克·施密特称它为“带着历史记录生活”，并告诫人们必须要让自己“以更加小心的方式说话，注意与他人的沟通，并慎重决定要将自己的哪些信息”提供给他人。言下之意很明确：避免暴露的方法就是不要去批评。然而，这固化了被监视者和监视者之间的信息权力差异，正如在边沁的圆形监狱中一样：因为史黛西·斯奈德不知道她的网页是否被监视、被谁监视，她就应当假设她的上司时时刻刻都在监视她，因此她应当更好地去表现，否则就会有被吊销教师资格证的风险。

至于公民与国家之间的关系，如果公民公开免责地批评政府的能力被削弱，那么我们就必须得承认这种信息权力差异着实让人不安。这就是为什么美国宪法第一修正案保护公民的言论自由而免受政府干涉的原因，它与《欧洲人权公约》以及其他许多国家宪法保障公民的言论自由非常相似。在缺乏这种保障的地方，公民往往都对公共事务毫无兴趣，极少出现健康的公共辩论，同时民主也遭到损害。纳粹德国、苏联及其卫星国中关于异见人士的历史，都提供了丰富而惨痛的例子。而在异见人士被允许的地方，他们要么保持缄默，要么就巧妙地进行伪装，使得当权者不把他们视为威胁。

史黛西并没有批评国家，也并不是政府阻止她获取教师资格证。只是因为她上传了一张名叫“喝醉的海盗”的照片到她的网页上，本意是想给她朋友们看的，但恰好这张照片被第三方的个人（她的上司）看到了，才最终被剥夺了教师资格。

也许有人会认为，这个事件会让她好好地吸取教训——她应当会成熟起来，并且不再把愚蠢的照片上传到互联网上。但是，我们怎样才能合理地预期哪些行为将不会给我们惹麻烦？我从未想过史黛西头戴海盗帽手拿杯子的照片会引来如此多的后患，竟然妨碍了她获取教师资格证。史黛西自己肯定也没想到。

不过，一旦我们意识到信息可以传递给任何人，那么我们将会表现得十分小心谨慎，而且如果有疑问我们也会选择审查自己，而不是选择承担不可预料的损害的风险。当然，这也并不新奇。然而在模拟时代，信息被划分得更细碎，更难以被意料之外的接收者获取到，同时也就更容易否定。另外，数字化记忆为了完全不同的目的，使更多的接收者可以获取更多信息，同时被认为保持了确凿无疑的精确性和客观性。

史黛西的例子强调了访问信息所衍生的信息权力，如何能够使一个

人开始考虑自我审查，那么正是数字化记忆的可访问性促进了这一行为。而安德鲁·费尔德玛的例子强调的则是数字化记忆的另一个特征。他没有意识到数字化记忆的持久性。他惹上麻烦并不是因为别人读了他的文章，而是因为别人在他写完多年以后读了他的文章。如果回头去看，史黛西在考虑到除了她的朋友外，还有谁可能会访问她的网页之后，她应该会自我审查一下；而安德鲁则本应该基于一个无法预料的未来去限制自己的写作。因此，安德鲁的例子让人更加不安。如果我们不得不想象多年以后，或许在未来几十年后，有人可能会如何解释和权衡我们的话语时，我们当初就会更加谨慎地阐述它们。如果说史黛西的例子是边沁圆形监狱空间版本的一部分，那么在其中，她不知道谁在监视她，但她必须假定她被所有人所监视；而安德鲁的故事说明的则是一个更为压抑的时间维度上的圆形监狱。

如果孩子们担心他们的直率之言可能会损害他们未来的职业时，那么，他们还会在学校报纸的网络版上坦率直言吗？如果我们担心我们想抗议的公司，在可预见的未来可能会拒绝跟我们做生意时，我们还会抗议这些公司的贪婪或者对环境的破坏吗？在民主国家里，个人既是公民又是客户。他们从事经济交易，同时又热衷于公共议题。有时，他们可能会发现他们自己所反对的，正是他们的交易伙伴所支持的。

在模拟时代，如果一个人没那么心直口快，他可以轻易地同时做到这两点：与自己心里所反对的交易伙伴做生意。

以汽车公司为例。一个人能够很容易地购买一辆汽车，与此同时支持汽车制造商所反对的高排放标准。假设我们的交易伙伴更加准确地知道我们的观点。他们还会跟我们做交易，为我们提供最优的价格吗？他们还有可能会雇佣我们吗？想一想就知道，他们不会，但却可能限制我们成为他们消费者的意愿，更别说让我成为在公共议题上直言的公民了。

在边沁的原著发表200年之后，法国哲学家米歇尔·福柯认为圆形监狱的观念已经蔓延到现代了。据福柯所说，不仅仅是监狱，而且其他等级森严的机构，比如军队、学校，或者工厂，都表现出一种无处不在的俯瞰和记录你生活的趋势。如果福柯今天仍然活着的话，他肯定将会表达这样的观点：数字化记忆作为一种全景控制的有效机制，不仅支持了对等级森严的机构和社会的控制，并且还会去寻求对他们自身的支持，从而巩固并加深现有的（不平等的）信息权力分配。

简而言之，这些说的都是与信息相关联的权力问题，以及它的相对性。当他人能够访问我们的信息了（尤其当我们并没有批准或者甚至不知道的时候），我们就失去了对信息的控制力。由于数字化记忆的可访

问性和持久性，信息权力不仅从个人手里转移到一些知名的交易方手里，而且还转移到了其他不为人知的机构手里。这固化并加深了现已存在的信息富民和信息贫民之间的权力差异，而且甚至可能会拒绝给后者提供他们自己关于他们过去的概念。同样有问题的是，它通过全景控制制造了一种自我审查的气候，这种气候不仅仅在当下，而且直到遥远的未来，都一直限制着健康而开放的辩论——这种辩论正是民主政府的标志。

永久的过去VS被忽视的现在：时间的威胁

我对于数字化记忆的第一个担忧是基于由此带来的信息的控制和必然导致的个人能力的减弱，第二个观点主要集中在数字化记忆如何让时间的作用失效，继而威胁到我们进行理智决策的能力。

为了理解人类的决策如何能够被数字化记忆所影响，让我们考虑一下下面的情景：

有一天晚上珍妮在家，她收到了她的朋友约翰发来的一封电子邮件。她认识约翰20年了，但是近5年她可能只见过他十几次。她仍然记得几个月前，他们最后一次见面时温馨的场面。他们的沟通方式是那种只有老朋友之间才有的愉快交谈。

现在，约翰发来电子邮件说他将要出席一个会议，而珍妮也将要在这个会议上发言，所以他想问问珍妮到时是否有时间一起喝杯咖啡。珍妮很高兴。一个想法闯进她的脑海：为什么不在几年前发现的一家小咖啡馆见面呢？它的名字叫什么来着？

然后，珍妮搜索她的邮件文件夹，寻找她和约翰互写的电子邮件，希望能找到那家咖啡馆的名字。几秒钟之内，显示出了好多邮件，时间几乎跨越了十年，邮件按日期整齐地排列着，最早的邮件显示在最顶端。她迅速地浏览了这些邮件。

随着邮件在眼前一闪而过，勾起她无限回忆：他们周末去海滨兜风时约翰驾驶的时髦汽车，他那丑陋的山羊胡子，以及她最终如何成功地说服约翰剪掉了它。多么有趣的时光啊。随后她偶然发现了一封完全不同的邮件；邮件很短，但是很伤感情，因为在这封邮件中，她在指责这位老朋友的公然背叛。珍妮情不自禁地读完了这封邮件，然后她又往下读了几封邮件。她看到了约翰愤怒的回复。此后的一年之中，他们不再有邮件往来，再后来也只是偶尔有零星的几封邮件，而且邮件的语气既客气又冷漠。她重读了这几封往来的邮件；她看到了当时自己言辞之中那种被背叛的感觉，他否认背叛时言辞之中的傲慢，电子邮件中那种死一样的沉寂让她犹豫不决，不确定曾经到底发生了什么，也不知道这次

争执到底是如何结束的。但是不管曾经发生了什么，她脑海中关于老朋友约翰的美好形象，突然以公然背叛她的形象出现在她的面前。他怎么能这样呢？珍妮甚至在想，她怎么会在几分钟之前还想跟他喝咖啡呢？

并非珍妮想要看到那些邮件，而是数字化记忆再现了她原本已经记不起来的事情，从而破坏了她对约翰良好的印象。如今，珍妮已经很难从脑海中擦掉那些不愉快的邮件。珍妮的这种反应是人之常情。这也是人类的记忆工作的方式。随着时间的推移，后来愉快的经历，让珍妮在记忆中完全擦除了对约翰的负面印象。为了获得一些信息，比如那个咖啡馆的名字，她去搜索了她的数字化记忆，却让她再次看到了多年前那些愤怒的邮件往来。她理性的头脑想要忽略那些陈旧的争吵。过去的事就让它过去吧！她与约翰的冲突发生在多年以前，而且即便约翰表现得比较不负责任，她显然也已经原谅了他，甚至都到了已经忘记了那次冲突的程度。但是，当她读到邮件中的那些愤怒之辞，那些负面记忆被触发了。

正是这些外部刺激，帮我们重新激活了我们原已淡忘了的与负面记忆的联系。慢慢地，当时与约翰争吵的一些不愉快元素重新返回了她的记忆。不过，她决定不理会这些负面情绪，依然到时去见约翰。但是，她真的能完全换一种新的心情跟他打招呼，而不露一丝愤怒的痕迹吗？就在一个小时前，她还会毫不犹豫地说“我可以做到”。但是现在，她不再那么确定了。

珍妮和约翰的故事是虚构的，但是我所描述的我们如何记忆、如何理解我们的过去的这种方式，则深深地根植在人类的行为之中。人类记忆的这种方式的两个特性，让这个图景变得复杂起来。

第一个特性与记忆如何消逝有关。我们都知道，我们对近期发生的事情的记忆比对更早发生的事情的记忆更清晰。但是，这只是一个粗略的经验法则。更接近真相的是，对于那些经常回想的事情，我们记得更牢固。记忆特有的这种行为方式，使得我们将来能够提取记忆片段的可能性更大了。

另外，那些我们更少去回忆的事情，记忆消逝得也越快，甚至都到了我们根本难以想起的程度。类似地，对于那些与我们目前所处背景拥有共鸣的事情，我们也拥有更好的记忆。如果我们现在很快乐，我们将会回忆起快乐；如果我们现在很沮丧，我们也就更难想起曾经的美好时光。而且，与那些和我们现在的信仰相冲突的记忆相比，我们更容易回忆起那些与我们现在的信仰相符的记忆。我们可能仍然能够回忆起，那些最初我们无法想起的、与当下相冲突的记忆，但是这可能需要一些广泛而持久的刺激，就像珍妮阅读她跟约翰之间过去的邮件那样。总而言

之，这指出了人类理解他们过去的方式中所固有的限制，而且如果我们面临着数字化记忆提供的线索与我们当下的直观回忆之间的紧张状态，我们将会非常挣扎该如何恰当地处理这些不同的记忆。

我们人类记忆方式的第二个特性是，我们的大脑无法处理被提取的所有类型相似的线索。从1978年到1984年，荷兰心理学家维勒姆·瓦格纳（Willem Wagenaar）保存了关于他这段生活的详细日记。每天他都把当天发生的最重要的事件记录下来，包括事件在哪里发生，发生时有谁在场，但是在保持记录的这些年中，他从不允许自己去看他之前写的日记——他不想让日记中的内容唤醒任何内部记忆。然后，他会试图通过使用日记中的部分线索，随机地回忆之前被记录下来的事件。

比如，瓦格纳会随机地阅读某一天发生的事情，然后试图去回忆事情在哪里发生、事情发生时都有谁在场，以及事情发生的具体日期是哪一天。正如所预期的那样，这些线索在他回忆那些事情时提供了极大的帮助，而如果没有这些线索他想记住这些事情将要困难许多。但是，出乎意料的是，不同类型的线索的作用并不同样有效。有一个明显的分水岭，使用“发生了什么”（what）去回忆“谁在场”（who）、“在哪里发生”（where）以及“什么时候发生”（when）效果最好，其次几乎同样有效的是使用“在哪里发生”和“谁在场”作为回忆的线索。

相较而言，“什么时候发生”这一线索几乎在所有情况下都没有作用——知道某个特定的日期，几乎无法帮助瓦格纳回忆起当天发生的重要事件。瓦格纳的结果反映出了，人类基于过去的时间挑选并提取记忆时存在极大的困难。

当然，一个确切的日期（比如，1971年4月30日）或许是一个太过人为化的时间标记，以至于人们很难将它与事件联系起来。既然很难回忆起特定的一天，瓦格纳就认为也许使用相关的时间点回忆事件时可能会更容易。为了验证这个想法，他每天写日记时将会记录两件事情，而不是一件事，他这样持续了几百天并且对这几百天作了标记。他希望能够将一件事作为回忆另一件事的线索——如果我们的记忆不是按照顺序来排列的话，也许我们记住的事情至少与前一刻或后一刻发生的事情存在一些相关的关联。但是正如他所发现的那样，事实上并不是这样。人类不仅基于具体的时间线索去回忆事情存在极大的困难，而且通过相关的时间线索去回忆很可能也存在困难。坦率地说，时间对于人类而言，是一个非常难以掌握的记忆维度。我们所有人都熟悉这一点。人们常常会谈到将一系列事情“串起来”，不是因为它很容易（如果容易的话，我们就不会这么说了），恰恰是因为它很难。

取舍之道

究竟为什么我们在脑海中通过时间顺序对事件进行排序如此困难，就像珍妮那样，到目前为止还不清楚。理解我们这个看似缺陷的问题，一个可能的关键是人类的遗忘。通过让关于我们过去事件的记忆随时间渐渐消逝，我们已经在适当的位置拥有了一个完美运行的机制，将过去的事件串在一个时间轴上。

因为生物性的遗忘就存在于我们人类的生理当中，通过人类几千年的进化，我们都未能发展出一种替代的认知能力，去正确地评估我们过去发生的事件。这样一种解释与通常情况下进化的运作方式相一致，进化就是一个实际而保守的过程。与那些更复杂的替代方法（也可能更容易出错）相比，那些经过试验和测试、已经能够良好运作的方法更受青睐。对我们而言，可能进化留给我们的就是生物性的遗忘，而不是详细感知理解过去时间的能力。

随着我们通过使用数字化记忆破坏了生物性的遗忘，使得我们自己更容易受到犹豫不决或者判断失策的影响，就像珍妮那样。这是数字化记忆的诅咒。它远远超出了信息权力转移的范围，已经涉及了我们人类及时行事能力的核心。

在博尔赫斯的短篇小说中，富内斯无法自抑地记住了他生活中每一个时刻的每一个细节。由于富内斯没有遗忘的能力，于是他就只能永远陷在他的记忆中，无法思考。博尔赫斯曾写道，“思考就是去忽略（或者遗忘）差异，思考就是去概括，去抽象。但是由于那场意外，富内斯不幸地陷入了‘只见树木，不见森林’的境地。在伊雷内奥·富内斯丰富的世界里，除了细枝末节别无他物。”对于博尔赫斯而言，完美的记忆力通过永不消褪的繁杂信息威胁着要折磨它的受害者，使其无法拥有清晰的抽象思维，因此就像《追忆逝水年华》中的马塞尔·普鲁斯特一样，将那些受折磨的人囚禁在自己的记忆之中——而正是遗忘这一人类独有的特性，才让我们得以从细碎的细节中抽身，把握住世界的一般规律。

如果我们没有遗忘的能力，那么每当面临抉择时我们往往都会回忆起我们过去所有的选择，这可能将会使我们犹豫不决，就像AJ。那个拥有近乎完美记忆力的女人这样告诉我们：

我记得好的事情，这让我感觉很舒服，但是我也记得那些不好的事情，以及每一次糟糕的选择。而且事实上，我无法让自己停下来休息。我所面对的都是这些岔路口，都是我不得不做出选择的时刻，而且哪怕过去了十年，我仍然会因为十年前的选择而责备自己。在许多事情上，我都不肯原谅我自己。你的记忆之所以存在是为了保护你。而我感觉我的记忆并没有保护我。大多数人都说我拥有一项天赋，但是我只觉得它是负担。

随后，在她最近刚出版的自传中，AJ写道："尽管人们往往把遗忘当成一种折磨，并且为由于年龄增长失去了如此多的记忆而感到困扰。但是我已经渐渐明白，能够遗忘我们生活中的许多部分，才具有真正的价值。"

从这种意义上看，遗忘并不是令人困扰的缺陷，而是一种足以救命的优势。当我们忘记了过去的时候，我们便重新获得了去概括、去概念化的自由，以及最重要的行动的自由。以珍妮为例，在阅读那些过去的电子邮件之前，她将约翰视为朋友。她心里早已忘记了过去所发生的冲突，这正是因为那些记忆已经不再重要了，因为那些记忆的价值已经被后来与它们立场相反的事件取代了。一种相对粗略的机制，也就是生物性的遗忘，使得珍妮可以假设约翰和她一直是朋友。另外，数字化记忆却将已经被遗忘的信息又带回来了，使得珍妮在做决定时变得很矛盾，失去了她应有的果断，并且被陷入可能做出错误选择的境地。此外，遗忘可能对学习过程很有帮助。正如组织学习专家威廉·斯塔巴克（William Starbuck）所说，学习一些全新的东西要求人们首先"忘却"陈旧过时的东西。生物性的遗忘正是这样一种极为简单又优雅的忘却方式。相较而言，数字化记忆能够如此历历在目地保存我们对已有知识的记忆，所以抑制了我们学习的能力。

因此，随着我们通过数字化记忆拓展我们对外部存储器的使用，我们正以多种方式危害人类的推理。其中的三种危害我已在前文中提到过。

第一，外部记忆可能作为记忆的线索，使得我们回忆起那些我们原以为已经忘记了的事件。如果人类的遗忘至少是部分基于相关性过滤信息的一种建设性过程，那么数字引发的我们对我们已经"忘记了"的事件的回忆，可能会破坏人类的推理。第二，广泛的数字化记忆可能会加剧，人类将往事按照适当的时间顺序进行排列的困难。第三，数字化记忆可能会使我们面临太多的往事，从而妨碍我们及时决策和及时行事的能力（与AJ和富内斯的经历相似），以及学习的能力。

第四个危险在于，当面临数字化记忆与人类对往事的回忆相矛盾的情况时，我们可能会失去对自己记忆的信任。就像第2章中描述的那样，其实人类的记忆相当准确，只是我们无法使存储的信息不受未来的影响。回忆不像从书架上拿下一本书，掸掉一些灰尘之后，包含的信息与我们当初把它放上书架时完全一模一样。

取舍之道

回忆是从记忆的资料库中提取信息，而记忆是不断变化的，而且在被重建时混入了后来的经历、喜好和偏见。尽管记忆已经被改写过很多

次，人类仍然信任他们的记忆，从而信任他们自己所记住的过去。有时，这会导致一种令人不安的情形，在这种情形下两个人对过去的同一件事拥有截然不同的两种回忆。

我们发现这种情形非常让人烦恼，而大脑往往会通过关注记忆中相同的那部分，而非记忆中那些不同的部分，来帮助我们克服这一烦恼。在数字化记忆偏离我们自己的记忆时，我们所记住的过去中的这种分歧将会发生得更加频繁，也显得更加明了。在这种情况下该信任谁？是信任我们的大脑还是外部记忆？逻辑告诉我们应该信任后者，但直觉却在诱使我们相信自己。数字化记忆无情地揭露出事实信息与我们独有的回忆之间的差异，在这个过程中我们可能会失去对自己所记住的往事的信任。

当然，这未必全是坏事。正如我所说，人类的记忆并不完美，它经常让我们失望。在这些情况下，数字化记忆可能会大有益处，因为当记忆差异出现时我们可能会迫使自己去信任数字化记忆而不是自己的记忆。信任自己所记住的往事，将会被信任数字化记忆所体现的往事取代。尽管需要良好的训练，但是数字化记忆具有融合各人记忆的优点——我们所信任的往事与别人所信任的往事是相同的。不过，更多地信任数字化记忆而非自己的回忆，使我们暴露在另一种挑战面前：假设外部记忆（尤其是数字化记忆）本身并非不可改变，而是可以事后被修改，从而不一定能准确地反映某件往事，那么将会怎么样呢？如果我们所有人都信任同一个信息源，那么我们就都会同等程度地受到这种可变性的影响。

对一个社会的过去进行控制是独裁政权的标志——即便必须伪造公文，也依然如此。比如在苏联，画家被命令将失宠的革命同志从公共的合影中清除掉，从而从共享的外部记忆中擦除他们出席过，甚至是他们存在过的证据。乔治·奥威尔写于半个多世纪前的著作《1984》令人震惊地详述了，当一个社会的过去不再受个人的控制而是被中央政府的执政者所蹂躏时，这个社会会发生什么。在这样的社会中，奥威尔解释道，“历史被不断地改写。真理部（Ministry of Truth）日复一日地伪造过去，这对维持政权稳定的必要性不亚于友爱部（Ministry of Love）所进行的镇压和侦察行动。过去，原本是记录和记忆达成一致的那部分事实。由于党完全控制了所有的记录，同样也完全控制了所有社会成员的头脑，所以党想要过去什么样，过去就是什么样。还有另一个结果就是，尽管过去可以被改变，但是在任何具体的实例中它都从未真正被改变过。因为每当过去被改造成此刻所需要的样子，那么这个新的样子就是过去，而其他样子的过去就从未存在过。”如果我们放弃对自己所记

忆中过去的信任，转而信任数字化记忆，那么独裁政权就再也不必控制我们的头脑。因为控制我们集体过去的外部记忆就足够了。这样一来，奥威尔的反乌托邦愿景实现的可能性也就更大了。

改变外部记忆不再需要经验丰富的苏联画家，或者像奥威尔的《1984》里那样伶牙俐齿的审查员。维基百科，一部成千上万用户经常访问且令人印象深刻的线上百科全书，就可以被任何有此意愿的人所修改。很多错误被很快发现，但是正如一份越来越长的清单所展示的那样，目前为止并非所有的错误都被发现了。随着图片处理软件如Adobe的Photoshop被广泛使用，即便是最随意的用户也可以修改一张数码照片，让原本不笑的脸变成笑脸，或者润饰酒窝和皱纹。事实上，现在已有家庭小作坊可以提供这样的服务，将前任配偶从节日照片或集体照中擦除。

因此，在数字化记忆的世界里，不再只有“老大哥”才可以改变过去——现在每个人都可以。

诚然，伪造并不是数字时代的发明。它已经伴随了我们上千年，从“君士坦丁赠礼”（一本伪造的八世纪罗马诏书）到米开朗基罗的丘比特（一尊他有意使其看起来像古董的雕塑）再到希特勒的日记。

但两个时代的伪造有所区别。第一个区别是：我们熟悉伪造的模拟记忆，会谨慎地去处理它，仔细检查它，并据此决定是否值得信任。我们觉得数字化记忆更可信，在某种程度上正是因为我们学会的鉴别模拟记忆的技巧，比如，检查文件的外观，以及它是否被篡改，这无法用来鉴别数字时代的伪造。也许随着时间的推移，我们能慢慢进行调整，并以适当程度的谨慎和鉴别伪造的技巧去处理数字化记忆。然而，在这期间，数字化记忆那种表面上的可靠性可能只会进一步蒙蔽我们。

第二个区别是由数字化记忆表面上的全面性造成的。在模拟世界，我们理解的外部记忆，一张手写的便条、一张照片、一个文档，只是组成过去某个事件的一个元素。要想尽可能完整地掌握一个事件，我们得接受一点，那就是需要以各种不同的方式搜寻不同的信息源，而搜寻将会导致像不匹配的拼图那样无法完美地将各种信息拼接在一起。与我们的记忆方式类似，我们并不是简单地回忆一个事件，而是重构了这个事件。要想重构好它需要投入大量的时间和精力，历史学家可以证明这一点。数字化记忆则完全不同，因为在我们的搜索请求之后，关于过去某个事件一整版的信息片段就会在几秒钟内呈现在我们面前。出于对搜索结果的快速和全面性的敬畏，我们假设搜索结果呈现给我们的，实际上就是一幅完整的拼图，而我们需要做的只是将它拼接在一起。这类似于，当我告诉我的一些学生他们的研究并不完整时他们对我的回

应："但是，这是我通过Google能找到的所有东西了！"

但是，数字化记忆产生的并不是完整的图景，充其量只是数字存储器捕捉到的关于它的那些片段。以我们假设的情景为例，假如珍妮和约翰针对他们的争吵进行了一次动情而诚恳的面对面长谈，那么将会怎么样呢？如果最终他们相互拥抱并开始重归于好，那么又会怎么样呢？这将会是珍妮评估约翰行为时非常重要的信息。但是，如果那些争吵是"当面"发生的而不是通过电子邮件，那么珍妮将无法从数字记忆中提取它们。在没有那些重要信息的情况下重构过去，会导致非常不完整的图景。当然，我们用来进行决策的信息几乎总是不完整的——这是不可避免的。但是，在模拟世界，信息片段被随机地丢失了。有了数字化记忆，那些不是以数字形式被捕捉的信息，以及未被送入数字化记忆的信息都会受到更多的偏见，而被排除在外。这是一种系统性偏差，它的一个问题是它不仅歪曲了我们对事件的理解，而且还会让我们被技术所蒙蔽。总之，由于数字化记忆只会详述数字化的信息，人类像珍妮一样信任数字化记忆可能会发现，这比仅仅只依赖自己的记忆还要糟糕，使得人们倾向于忘记那些不再重要或相关的信息。

我开始有这样的感觉，数字化记忆使时间失去了作用。通过触发我们对自己原以为早已忘记的事件的回忆，数字化记忆有能力用相矛盾的记忆让我们困惑，并可能会影响我们的判断。我们可能会做出错误的决定，或者犹豫不决从而无法采取行动。我们可能会不再信任自己的记忆，进而不信任自己所记住的过去，并且用人造的过去而非客观的过去替代它。这不是我们的过去，也不是其他任何人的过去；它是一种从数字化记忆存储的有限信息中，重构出的人造过去，一种缺乏时间线索、完全扭曲的拼图，而且可能会同时被它所包含和不包含的东西所篡改。我担心，通过否认我们自己所记住的过去，我们可能会损害自己的判断能力——而且比想象中更严重。

除此之外，还有另一个维度显示了数字化记忆如何让时间失去了作用。前文中我曾提到，数字档案和人们在搜索引擎中查询所得到的结果都很古怪，因为这些结果仅限于以数字格式可获取到的信息，从而忽略了没有存储到数字化记忆中的事实——极有可能，还是很重要的事实。同样古怪的是，多年，也许是几十年的档案都被扔在一起：一个人在电脑中找出十年前的贷款记录，可能会发现旁边就是她六年后发表的学术论文的网络链接，以及上个星期她才上传到Flickr网站的照片。

这样产生的拼贴画加上所有遗失的，同时可能也是误导性的信息，并不能准确地反映这个人。使用更为全面的数字信息源，也没有太大的帮助。结果产生的数字事实拼贴画，仍然会问题重重。这就好像把一盒

关于你的、没有分类的照片扔到桌上，以为仅凭足够努力的观察，就可以获得关于照片上的人如今全面而准确的感觉。

也许你认为上面的例子太过牵强，那么看看Flickr网站吧，看看它是如何使时间跨度非常大的照片可以一起获取的。这些数字拼贴画与无数的关于我们的信息相结合，每一个（充其量）只在我们过去的某一点上是有效的。但是，正如它呈现给我们的一样，它呈现给我们的是时间被消除了的信息，同时也是一幅拼贴画。在这样的数字拼贴画中，改变只有从两个矛盾事实相冲突的角度才能看到，而无法从一个演化过程随着时间变化的角度看到。

有些人担心，数字拼贴画类似于关于我们瞬间而全面的快照，就像一张照片一样，可以被全世界所获取。实际上，数字拼贴画比这更令人不安。它们不是像一张照片，而是几百张，也许是几千张关于我们生活的快照（而且彼此相互重叠），但是却没有时间的线索。我们如何能够以这样的方式抓住某个人的感觉呢？我们何以希望去理解多年来一个人怎样发展，怎样调整他的观点，怎样适应（不断变化的）环境？当呈现给我们的所有信息都只是没有时间线索、放在一起的个人事实拼贴画时，我们如何能够假装理解一个人如今的状况呢？他的价值观怎样，他的想法怎样，他的性格发生了什么变化？

也许，数字化记忆的支持者将会反驳说，我们可以通过适当的数字过滤器做到这一点。在这个过滤器中，通过视觉呈现，时间能够重新被引入，就像用不同颜色编码年份，或者分类事件一样。然而，这恐怕仍然无法完全解决这个问题。因为，我担心，忽略时间而将不同年代的历史合并在一起带来的真正瓶颈，不是数字化记忆，而是人类的理解能力。即便给我们呈现一个事实档案，通过日期整齐地分类，但是在心里我们仍然难以将这些事件以一个正确的时序串起来，仍然难以随着时间的变化恰当地评价这些事实。就像珍妮遇到的苦恼。从个人记忆（remembering）的角度来看，数字化记忆阻碍了判断。然而，从个人已记住（remembered）的角度来看，它否认了发展，并且拒绝承认所有人类时时刻刻都在变化。通过永久地“召回”我们的每一次小错和大错，数字化记忆拒绝了我们人类从中学习、成长和发展的能力。

如果人类的行为永远不会被遗忘，那么人们几乎没有必要去提升他们自己并寻求改变。在一个历史无处不在的世界里，可能也几乎没什么动机去努力奋斗逃离自己原本较低的社会阶层，也没什么动力从自己既有的模子中摆脱出来，而这些都是现代文明社会的基本构成元素。当然，即便没有动机，人类作为一类生物也将会在数字世界里继续改变——我们的身体将会变老，同时也将会修正我们的观点。但是，我们的

数字表征将永远把我们拴在我们过去所有的行为上，实际上使得我们不可能从中逃脱。没有了某种形式的遗忘，原谅则成为一件非常困难的事情。或者正如T.S.艾略特所写的那样，“如果所有的时间都永远是现在，那么所有的时间也都不能够得到拯救。”

记忆阻碍了改变。这对所有的记忆而言都是对的。然而，在模拟时代，记忆仍然很昂贵，因为广泛、及时又实惠地获取它在很大程度上只是一种幻象。我们谨慎地使用外部存储器，而不是随随便便使用，或者一直都使用。存储器被爱惜而审慎地使用，因为这是一种非常有价值的资源，它把我们训练得在决定是否使用存储器时，就像对待精致昂贵的香料一样慎重。而在今天，数字化记忆无处不在，如此廉价，而且看似“很有价值”——这得归功于数字化记忆的可访问性、持久性和全面性，以至于我们经不住诱惑不停地所用它。以这种不加区分的方式使用，数字化记忆不仅使过分拥有它的人判断变得迟缓，而且还否认那些记得发展的时间和空间的人的存在。

取舍之道

关于时间，还有一个终极难题。正如我之前所说的那样，数字化记忆可能会促使我们停止对我们人类记忆的信任，转而偏向数字化记忆。这样，数字化记忆就被抬高到重建过去的首要来源的位置上：便于使用又范围广泛，而且看起来比健忘的人类大脑那容易犯错的记忆更加准确和客观。但是，与原子相比较，信息位更具有延展性。正如我提到的那样，它们很容易被改变，从而使历史被改变。当人们意识到过去不再值得信赖时，会发生什么情况呢？或许这个问题的答案，会与某人问人们将如何与永不会遗忘的数字化记忆一起生活这个问题的答案类似，因此否认了改变的作用。如果我们的过去，数字化记忆构建的过去，不再值得信赖，并且不再有用，我们可能会忽视它。

在最好的世界中，这将使得我们回到人类的遗忘（与记忆），或许还伴随着对数字化记忆罕见但有价值的使用，如果这样的话我将会欢呼。但是事情可能发展得不会有那么好。我们可能进入了某个时代，在其中作为对太多记忆的一种回应，它将会拥有与过去过为严格和无法原谅的链接——一些人可能会走极端，一起忽视过去和现在，决定仅仅活在当下。

这听起来可能比它的实际情况更牵强附会。以短信文化为例。正如互联网企业家伊藤穰一（Joi Ito）提醒我的那样，大多数来回发送的短信（很像某个人的facebook状态）并不是打算谈论过去或者未来的什么事情，而仅仅只是关于现在的事情：这是我所在的地方，这是我做的事情。这些短信是特意设置成很短暂的，抓住飞逝的当下，而非遥远的时

刻。至少在某种程度上，这可能是数字一代应对数字化记忆的一种方式。但是，这种对此时此地的过分强调可能将婴儿和洗澡水一起倒掉了，导致可以理解但不健康地徘徊在另一个极端里——从一种普遍传播的过去到一种完全被忽略的现在。而这些极端情况没有一种是有用的。相反，我们需要认识到，人类的存在同时拥有记忆和遗忘的能力。

取舍之道

在本章中，我已经指出了随着我们无法遗忘，我们所面对的困境。在“信息权力”那部分中，我描述了数字化记忆如何加深了信息富民与信息贫民之间已经存在的鸿沟，如何进一步增强前者对后者的损害。我解释了完整的数字化记忆会如何扼杀社会辩论。在信息权力与时间的交汇处，永久的记忆创造了空间和时间圆形监狱的幽灵，在其中，每个人都可能不停地被诱使去进行自我审查。或许最为重要的是，完整的数字化记忆摧毁了历史，损害了我们的判断和我们及时行事的能力。它拒绝了人类进化、发展和学习的机会，使得我们无助地徘徊在两个同样让人不安的选择之间：一种选择是永久的过去，而另一种选择则是被忽略的现在。

05 来一场“互联网遗忘”运动：应对数字化记忆与信息安全的6大对策

数字化记忆仿佛是一个诅咒，人类对它愈发强烈的依赖阻碍了我们从中学习、成长和发展的能力。而信息隐私权的维护不仅应在当下，还应在未来。如果有一天隐私权被废除，信息处理者们坐在堆积如山的个人信息中为所欲为，那么信息隐私权还有什么意义？

如果我们相信世界上某一种系统中的信息是完整的，那就是自欺欺人。

只要数字化记忆仍然存在明显的不完整，那么它就不会比人类的遗忘行为好到哪里，甚至还要更糟。

为了完全理解一个过去做出的决定，我们不得不潜入一片相关数字信息的汪洋大海。即使是完全语境化，可能也无法使我们全面地理解过去。

如果数字化记忆可能会造成如此严重的后果，那么我们该如何应对呢？在本章中，我将详细阐述六种可能的对策，旨在预防或减轻数字化记忆对信息权力与时间所构成的挑战。这六种对策不仅具有实质性差异，而且在可行性上也各有千秋。其中两种已经在相当多的社会中得到了应用。就这两种对策的效果而言，目前已经拥有相对较好的（甚至可以说是优质的）研究数据。其余对策虽已提出，但尚未得到广泛应用。因此，我对它们的评估难免思辨色彩更浓。

对策1：数字化节制

这种方法基于这样一个观点，即个人在应对数字化记忆一事中扮演着中心角色。其核心是：一旦人们认识到了放弃遗忘的潜在影响，他们将不再向他人提供个人信息，数字化记忆将因此不复存在——至少不会以我之前所说的那种全面且具有威胁性的形式存在。如果史黛西没有把照片上传到MySpace上与大家分享，她可能也不会惹上麻烦；如果安德鲁·费尔德玛没有将手稿发表在一份日后也能在网上获取的期刊上，他或许仍然能够不受约束地前往美国。如果人们能够对上传至互联网的个人信息有所节制，那么便无需担心长期存在的数字化记忆会造成的后果——既不必担心失去信息控制权，也不必担心暴露在“数字圆形监狱”或推理能力损害中。

取舍之道

解决方法看似简单明了：尽可能远离那些向他人透露个人信息的互动。为了实现这一行为转变，每个人需要接受教育，以了解自己在日常生活向他人提供了多少信息，这些信息的价值，以及借助数字化记忆将

相关信息存储和挪作他用的危险。数字化节制假设，一旦人们发现其个人信息有可能被滥用，他们就会对自己的行为做出调整。了解到这一点将会为变革播撒种子，导致民众运动，以节制向数字化记忆输入个人信息的行为，进而使数字化记忆陷于瘫痪，社会重新恢复遗忘。数字化节制带有某种反科技进步的勒德主义色彩。

然而，与某些主张全盘放弃科技的顽固派新勒德主义不同，数字化节制单纯地专注于通过个人选择来限制不可控信息的流通。

数字化节制建立在个人知情与偏好的基础上，不依赖规章制度和相应的执法，而是立足于个人决定，立足于个人选择遵循的行为规范，而非法律或其他强加于他们的外部约束。

这一方法的成功取决于一系列因素。首先，它通过影响人们的行为而起作用。这就需要让人们了解误用个人信息的危害性以及如何避免。说起来容易做起来难。多年以来，媒体对侵害信息隐私可能造成的消极后果一再进行过报道，但是我们却仍然没有看到公众态度发生显著改变。事实上，正如哈佛大学伯克曼中心的约翰·帕尔福瑞和乌尔斯·加瑟所说的那样，世界各地年轻的数字原住民更加包容，在同他人分享个人信息时也更加放心。在他们的世界中，数字化节制，以及数字化记忆可能存在的危险都没有得到广泛传播。

当然，这一情况会得到改观。随着关于数字化记忆的负面报道越来越常见，史黛西·施奈德和安德鲁·费尔德玛经历的广泛传播，更多人将会注意到这一点。我最初撰写的关于数字时代遗忘之好处的论文产生了令人惊喜、振奋人心的影响，至少对我而言。影响之一就是公众对于了解数字化记忆的潜在危害有了强烈而持久的兴趣。如果这能说明一些问题，或许想要在教育用户们如何限制个人信息泄露这一点上，做第一个吃螃蟹的人也并非不切实际。

然而，让人们了解数字化记忆及其潜在影响还只是第一步。如果在实际行为中没有后续的改观，它将是毫无用处的。要想使数字化节制行之有效，人们必须开始节制在互联网上发布个人信息的行为。他们还必须选择交易伙伴，根据其对个人信息要求的多寡，对亚马逊等电子商务网站或谷歌等搜索引擎加以选择。

问题在于，分享个人信息能够为用户提供价值，人们或许并不想放弃。作为对获取个人信息的交换，电商能够向顾客提供个性化交易体验，如亚马逊提供的图书推荐，谷歌提供的更多有针对性的搜索结果。人们难道会一致选择放弃这些好处，与数字化记忆的危害作斗争吗？这不仅仅涉及质量、效率和经济利益。整整一代人都是在利用互联网中成长起来的，不仅用它获取信息，而且分享信息。同史黛西·施奈德一

样，超过1亿人都在MySpace上拥有个人主页。而同安德鲁·费尔德玛一样，每三个美国年轻人中就有两个在互联网上与全世界分享内容。大多数年轻的互联网用户都乐于生活在一个各种信息紧密联系的世界里，在其中与他人分享自己的想法与经历，扩展机遇，并高效地进行交易。他们不大可能乐意放弃这一切并遵循数字化节制，除非分享信息会使他们暴露在某些现实的威胁中。要证明这一危害很难，至少在短期看来如此。数字化记忆的消极后果现在还很少对我们造成冲击，但它们还会再卷土重来，不断困扰我们。为了防止可能存在的未来危机而说服人们放弃当下的现实利益，这真有些因噎废食了。

或许这不失为一件好事。哈佛大学法学教授尤查·本科勒曾提出过极具说服力的论断，全球数百万互联网用户共同参与创造网络内容，从而产生了各种新颖而有益的信息生产形式，而这些形式在一个奉行数字化节制的世界里是不可能实现的。同时，要实现约翰·西利·布朗的数字拼装[1]。也将面临更大的困难。如果这些步骤意味着放弃尤查·本科勒所说的生产与分享的社会创新，那么或许数字化节制这一理念本身太过激进，太二元对立，它迫使我们成为数字时代的隐士，放弃数字时代的诸多便利。

如果数字化节制的核心理念，即人类应当在向互联网添加信息时更加谨慎，能够不那么极端，它或许可以行之有效。比如，人们可以想象出某种“温和版数字化节制”出现。与其指望个人保留绝大多数个人信息，不如要求人们在发布信息时更加谨慎。这一方法认为，在许多情况下，向他人提供信息能够创造价值，不仅能为集体，也能为个人创造价值。以此类推，温和版数字化节制的目标仍然是教育和提醒人们在分享个人信息前慎重选择。由于信息处理者们，包括从电子商务服务商到信息分享平台服务商，无法再让用户与其分享个人信息，他们就不得不调整商业活动并接受对数字化记忆的实质性约束。

这一“温和”的方法克服了之前数字化节制观念固有的极端绝对性。其效果取决于个人能够且愿意向信息供应商施加压力的大小。与数字化节制及其旗帜鲜明地拒绝分享个人信息不同，不幸的是，温和版数字化节制的优点——对灵活性和个人判断的依赖，也极可能是其软肋，原因有两点。

首先，个人在与信息处理服务商们谈判是否发布信息时完全各自为战。由于双方既存的权力不平衡，这种谈判在结构上就不利于个人。信息处理者们能够轻易地整合有关个人数字化节制偏好的信息，以影响谈判和单个用户的决策（比如通过更巧妙的消息提示）。与之相反，个人

只能在对他人决定毫不知情，以及完全不了解谈判对象会作何反应的情况下做出决定。这种固有的信息权差异，会使个人在坚持温和的数字化节制过程中遇到困难。

其次，个人陷入了两难境地。如果其他人都拒绝向某一既定服务提供个人信息，使得服务商在压力之下调整其对信息收集和存储的态度，个人也是可以从中获利的。同理，个人用户私下向服务商提供个人信息能够在短期内获利，比如，获得某项特殊服务。这样做有可能会显得像个“两面派”：一边在台面上广泛宣扬数字化节制，一边私底下大规模地背叛以获得信息分享的好处。

个人没有能力也不情愿放弃信息分享，这一事实将削弱温和版数字化节制反制数字化记忆的效果。然而，即便人们调整行为让自己更加慎重，更有选择性地与他人分享个人信息，数字化节制的效果仍然值得怀疑。信息很难追踪和限制。一旦我将信息传递给某人，我就必须信任那人会对该信息保密，但我却不再对其拥有直接控制权。所以，即便某人只向少数精心选择的交易方提供个人信息，仍然无法保障接收方会按约定使用个人信息。当这些信息从一个服务商传递到另一个后，即便不是完全不可能，也将很难找出是谁辜负了信任，并把信息泄露给了他人。没有哪种数字化节制能应对这样的泄露，因为节制的理念纯粹聚焦于个人发布信息的决定，而没有将随后服务商是否与第三方分享该信息的决定考虑在内。最后，即使我们在未来进行数字化节制，现在已经存在的海量个人信息将在未来数十年继续威胁我们。既然我们的个人信息已经存在，并且与他人分享了，为什么要停止泄露个人信息呢？

取舍之道

数字化节制初看之下十分吸引人，但我们一旦意识到很难让人们对自己渴望分享的事物加以约束后，这种方法的效果就大打折扣了。既想在停止数字化记忆与选择性公开之间找到平衡，又不得不看到个人信息一旦发布就很难限制。即便个人调整其信息分享行为方式，信息处理者们也缺乏动力来依样画葫芦。

这种执行起来的弱点能够以法律来解决，而非社会规范和个人不一定受用的营销刺激。这是信息隐私权的管辖范围。

对策2：保护信息隐私权

信息隐私权定义了一种可行且可接受的执行机制。这些权利形态各异，但享有共同的核心原则：向个人提供法律认可的个人信息权力，从而赋予他们维持信息控制的权利。这样看来，信息隐私权似乎是对数字化记忆侵蚀信息控制权的适当回应。信息隐私权最基本的形式是给予

个人选择是否分享信息的权利。如果某人不经他人同意而通过窥探获得其个人信息，他就触犯了法律，将要面临法律制裁。

最早提倡这些隐私权的知名人士是波士顿律师赛缪尔·沃伦和路易斯·布兰代斯（后来成为最高法院大法官的布兰代斯），他们于1890年在《哈佛法律评论》上发表了《隐私权》，提倡将侵犯隐私视作民事侵权处理。据说，他们是在地方报界刊登了对沃伦宴会的恶意报道后开始着手这一课题的。这可以解释他们希望处理的威胁为何如此有限：侵犯隐私，或按照信息时代的说法，未经个人同意而使用其个人信息。

这也揭示了信息隐私权面临的一个根本挑战。信息隐私权是否能有效保护个人对个人信息的控制取决于这一权利究竟是如何构建的。我可能同意邮局获悉我的新地址，但绝不愿它将其传递给第三方。我怎么能防止这一点？如果我对自己的个人信息享有简单的产权，但将信息转移给信息处理者就等于放弃了对该信息的任何控制。这就没有必要了。因此，信息隐私权需要一个精心构建的法律机制来实现，主要原因在于，个人在实践中极难确定信息泄露源和恰当的诉讼对象。

试想这样一种情况：某个人服用抗抑郁药的信息可能被医生、药剂师、保险公司、任何同样服用的人，甚至是朋友、配偶或知情的同事泄露。要想确认泄密者并采取有效措施几乎是不可能的。这也就是为什么全世界的司法体系都认为，信息隐私权应该将某人个人信息的所有后续使用者捆绑在一起，而不仅仅指个人委托的那一方。

然而，个人可能丧失信息控制权的危险，不仅体现在某个人信息的合法接收方在未征求个人同意的情况下将其传递给第三方；危险还存在于个人为了某一特殊目的而向某人提供个人信息，而该信息后来却被挪作他用的情况中。举例来说，如果银行为了明确获悉某顾客是否有贷款资格而收集和存储顾客的个人信息，顾客可能不愿银行在未获得其明确许可的情况下将该信息直接用于信用卡营销。这就是为什么更成熟的信息隐私权包含了目的限制原则——某一种法律约束，即个人信息接收方只能将该信息用于已许可的目的，而不能挪作他用。

信息隐私权也在范围上存在巨大差异。美国的1974年联邦隐私法案仅仅涉及联邦机构对个人信息的处理。联邦法律没有赋予美国人面对联邦政府之外其他任何机构的信息隐私一般权利。而欧洲的立法机关则较为大胆（或许归因于野蛮暴力的20世纪留下的历史阴影），不仅赋予了个人面对中央政府时的信息隐私权，而是面对所有公共或私人领域信息处理者时的信息隐私权。另外，从20世纪80年代中期的德国开始，欧洲人逐渐形成了一个范围更广的信息隐私权概念，将沃伦和布兰代斯的隐私权概念远远抛在后面。

如果说原信息隐私权更加专注于个人意愿，如今的信息隐私则被视作个人如何规划其社会活动的权利。

取舍之道

在一个互为联系的世界里，信息隐私权不可能再被理解为某种二元对立的许可权，不是绝对的是非，而必须以一种更微妙的方式重建，将许可与信息处理的特殊目的和条件相结合。它是一种从根本上决定个人如何参与，而非是否参与信息分享的权利。这一更加宽泛的概念常被称作信息自决权（informational self-determination），其目的是使个人对其个人信息每一阶段的用途拥有控制权。

这是一个成熟却复杂的体系。当信息隐私权只给予个人是否许可他人对其个人信息进行处理时，执行起来相对还算容易。信息处理者们只需证明其获得了每一个人的许可。而一个更宽泛的信息自决概念意味着个人许可应建立在某一特定目的以及信息处理场合之上。这一点迫使信息处理者们不仅要记录提供许可的人，还要记录因何种目的而许可。同样，这一点也适用于法律强制个人提供信息的情况（试想一下超速行驶被警察拦下后出示驾照的情况，或者破产流程中向法庭提交财务信息的情况）。

这样的结果就是，宽泛的信息隐私权与目的限制规则的有力结合。相比数字化节制，它有两点优势。首先，个人现在有了手段——法律系统，来实现其对信息的控制。其次，目的限制规则能够保障信息在未经个人许可的情况下不被重复利用和分享。然而不幸的是，这样的隐私保护法律机制也伴随着三大缺点：颁布起来难得出奇，至少在美国如此；使用率低得惊人，至少在欧洲如此；此外，还取决于未来社会是否愿意继续遵循当下的信息隐私原则。

赋予个人对其个人信息的控制权就必须限制他人的信息收集、存储、处理和散布，而这些人可能认为自己的活动受到某种基本权利的保护，比如言论自由，或至少是市场经济核心成分的基本经济自由。随之产生的紧张关系使得综合信息隐私的立法超出了公共领域的范围，并且使得某些对私人用途的狭隘限制在美国变得政治上不可行，至少目前为止是这样。在过去30年里，美国国会不时就更具综合性的信息隐私立法展开辩论，但并没有产生民众运动。政治学家或许能笼统地解释原因。他们表示，法律往往在受到某一界定明确、积极性高、经费充足群体的拥护，并且大多数公民能够容忍微不足道的损失时得以顺利颁布。

综合信息隐私权立法则刚好相反：大多数公民认识到能够从中获得少许好处（但因获利太少而缺乏干劲），而积极性高又经费充足的信息处理者群体则唯恐自己会失去很多。当然，考虑到可能存在的严重后果

（以及高损失），数字化记忆最终可能会触发一场规模够大的要求信息隐私的民众运动，迫使国会出台一部综合信息隐私法。然而，鉴于过去数十年的政治动态，我可不会把赌注押在这上面。

美国之所以没有通过相关联邦法规，以赋予个人广泛的信息隐私权，原因之一就是政治。另一个原因可能是对其效用持有严重怀疑。欧洲国家在赋予个人信息隐私权方面已积累了数十年经验，其中尤以德国最为突出。

20世纪90年代初，我曾请求所有德国数据保护专员给我发送一些利用信息隐私权使未经许可收集、存储或处理其个人信息的信息处理者付出经济代价的成功个案。他们的回应令我大吃一惊，他们竟然对此闻所未闻——居然一个案例也没有！诚然，他们不可能知晓所有案例，但这一结果仍然令人困惑。

为什么个人无法在法庭上维护其信息隐私权，可能的解释之一就是，法律维权大多面临风险，而且费时耗力。另外，要想在法庭上维护信息隐私权，个人就不得不暴露其身份，而他们或许并不想这么做。维权所得也太少，因为德国的责任法只允许实际损害获得赔偿（计算信息隐私案例中实际损害的难度是出了名的）。

那么，或许我们应该在维权的必要法律程序上做出变革，以支持信息隐私权。措施可能包括将提供证据的负担从个人转移到信息处理者身上，使刑事法庭而非民事法庭来执行维权，或通过降低诉讼费用来鼓励个人为其信息隐私索赔提出诉讼。确信这将带来重大影响后，欧洲立法机构修正了其信息隐私法规，以吸纳众多此类措施。这些修正已经出台了十几年，但我在欧洲各地信息隐私机构的非正式查询却显示，改变相对很小；个人诉讼信息处理者的案例仍然很少。当然，有可能信息处理者们全方面遵守信息隐私权，所以根本不必进行法律维权；或者维权行动发生在民事责任索赔的狭隘范畴之外。还有可能是维权措施执行不佳或没能有效传达给公众。然而，这样一成不变的结果肯定让大西洋两岸的信息隐私权提倡者深感忧虑：至少在欧洲，个人似乎不情愿投入足够的时间和精力来维护自己的信息隐私权。

信息隐私权背负的第三个固有缺点涉及信任的缺陷。信任是必要的，因为个人在与他人分享个人信息的同时也失去了对其直接的控制。个人可以相信信息处理者们不会误用其个人信息。这是数字化节制的观点。他们也可以信任某种社会制度，比如法律，来实现普遍尊重（即尊重信息隐私权）。信息隐私权的优势在于，个人在与他人分享个人信息后仍能保留某些合法控制。然而，为了保证这样的控制能够持久，维权不仅应在当下，还应在未来得到保障。如果一年之后信息隐私权被废

除，信息处理者们可以坐在成山的个人信息中对其为所欲为，信息隐私权还有什么意义呢？因此，某些国家已经给予了信息隐私权宪法保障，使其日后更难遭到废除。然而，以公然废除的方式来削弱信息隐私权的效用可能并无必要，因为这些权利依赖于政府执行它们的意愿。政府如果不再相信隐私，就不会大力加以维护，个人也就因此变得孤立无援。这听起来似乎十分抽象，也不大可能，但20世纪为我们提供了一个耐人寻味的例子。

20世纪30年代，荷兰政府实施了一次全面人口登记，包含了每个公民的姓名、出生日期、地址、宗教信仰和其他个人信息。这次登记被誉为对政府管理的促进和对福利策划的改进。后来纳粹德国入侵荷兰，并占有了这次人口登记的信息，无情地将数百万荷兰公民的个人信息用于确认、迫害、谋杀犹太人与吉普赛人。由于掌握了全面人口登记所包含的信息，纳粹得以确认、驱逐和杀害更多的荷兰籍犹太人（73%），比例远远高于在比利时（40%）、法国（25%）或其他欧洲国家的情况。甚至连荷兰境内的犹太难民也比荷兰籍犹太人的情况要好，因为难民（相对于犹太裔公民）能够免于被记录在人口登记中。纳粹借助人口登记实施迫害的另一族群——吉普赛人的境遇更糟。

这是个深刻的教训：荷兰人出于对政府的信赖而提供个人信息，完全没有预料到纳粹的入侵。我们或许因为生活在民主社会而感到安全，可那时的荷兰人也是如此。我们不知道未来会有什么在等着我们，不知道未来的政府是否会珍视我们给予的信任，保护我们的信息隐私权。

难以立法，效果可疑，未来没有保障，信息隐私权反制数字化记忆的总体适用性尚不明确。

或许，问题在于我们误解了信息隐私权，并且错误地将其生搬硬套于侵权等已经存在但并不适用的法律范畴，或者像信息自决权那样颇具新意但过于复杂的权利。对信息隐私权的另类构想或许能有效地保护我们免受综合数字化记忆的威胁。伯克利大学法学教授帕梅拉·萨缪尔森建议采用一种契约性方法，该方法利用商业秘密法的某些元素，可能更有成效。据她称，这有可能促成个人信息许可默认规范的出现，从而保证个人保留其对信息的控制（及权力）。这样做比僵硬的信息隐私权体系具有更灵活的优势，并且建立在已经认可的机制上，而非依赖复杂的新法律工具。相比关注于寻找实施信息隐私权最适宜的法律机制，纽约大学的海伦·尼森鲍姆教授提出，应当关注信息收集和传播发生的场合。在她看来，保护个人信息权的任务在于确保信息不会与其发生的场合相分离。尼森鲍姆的观点是目的限制原则的必然推论，然而，她的目标不在于审视信息的使用目的，而是将其使用场合考虑进来。这样可以

防止我在前面提到的"去关联化"，即便是在信息使用的总体目的保持不变的情况下。乔治·华盛顿大学法学教授丹尼尔·沙勒夫提出了另一种方法，他认为，信息隐私权应当关注信息的名誉方面。这一观点认为，危险不在于个人失去对信息的控制，而是个人的名誉因为该信息而败坏。传统的信息隐私权以及萨缪尔森、尼森鲍姆等人的提议更专注于信息如何（并且在什么条件下）从个人流向信息处理者，而沙勒夫则更关注信息的（消极）影响。

或许我们误解了产权的固有局限性。这正是伯克利大学法学教授保罗·施瓦茨最近令人印象深刻的主张。他坚持认为，产权概念拥有足够的灵活性和适应性，因此适用于信息隐私。这一信息产权化模型还致力于完全保障信息隐私，其关键在于将信息产权理解为某种能够通过法律系统来形成的利益捆绑。这使法律得以限制某些种类的交易，尤其是未经个人许可继续使用或转移信息的行为。

以上这些以及其他各种改革信息隐私的提议都有待考验。然而，它们至少已经实现了部分目的。它们详细阐述了某些现存信息隐私权体系的不足之处，并证明克服了某些缺陷的替代性权利框架是可能存在的——至少在理论上可能。未来几年，我们很有可能将会看到更具新意，更加精细的替代方案得到提出，甚至是立法。这是一片值得拭目以待的空白——因为个人可能会从改进的信息隐私权中收获颇多。

取舍之道

关于隐私权，这些提议（以及更传统的信息隐私权）的效果不仅依赖法律系统的执行能力，还依赖个人对触犯者采取行动的意愿。可以肯定的是，法律相比个人自制要更有优势，因为法律有能力限制人类行为，尤其是信息处理者的行为。由此看来，保护个人信息隐私权可能仍将是进展缓慢而代价昂贵的。因而，如果维权流程能够得到简化，或许个人控制（并保护）其信息的主意可以行得通。这是技术性替代方案相对于规范行为的传统法律机制的潜力所在。

对策3：建设数字隐私权基础设施

个人并不是唯一强烈感受到失去信息控制的群体。知识产权所有者在过去数年里遭受了类似地打击，原因是技术发展无情地淘汰了过时的商业模式，改变了用户的行为。几乎他们所有的内容都是以数字格式的形式存在。便宜的存储，全球宽带网络和价格低廉的网络月费，加上简单易用的文件分享软件共同造就了知识产权所有者们眼中的最糟情况：大多数年轻人免费互相分享内容，全然无视现存的版权法。版权所有者的回应一直是多管齐下。数万起针对个人文件分享者（不时还有无辜的

局外人）的诉讼出现在头版头条。然而，将版权维护嵌入技术之中的初步行动则可以说意义更加重大，且无处不在。

数字版权管理（DRM）的原则十分简单：在版权的范畴中，信息是指音乐、电影、游戏、数字图书，而在遗忘的范畴中，信息可以是任何个人信息——信息与关于使用者与使用方式的元信息相匹配。媒体播放器检查这些元信息，并拒绝播放未获得适当授权的信息内容。为防止这些内容被未经授权的设备播放或复制，内容和元信息常常被加密，需要一个只有获得授权的设备才能“识别”的特殊密匙。因此，同信息隐私权不同，DRM维权几乎完全建立在技术上。

初级DRM系统已经与我们相伴多年。许多商业DVD都被加密并与某一特定地理区域相关联；人们无法播放这样的光盘，除非拥有密匙或使用一台来自相同区域的播放器。类似地系统也被运用于视频游戏和游戏主机上，例如索尼的PlayStation或微软的Xbox，以及从DAT到MiniDisc等各种数字录音系统。数年来，所有通过苹果的iTunes音乐商店购买的音乐都内嵌有DRM元信息，限定了该歌曲可以下载到哪一个iPod上（个人必须通过电脑“注册”），或能够将其“烧录”进CD的次数（只能烧录七次，之后必须变更歌曲“播放列表”）。在个人电脑上播放媒体文件可以通过微软的Media Player（结合受保护的WMA格式）和RealNetworks的RealPlayer（与RealPlayer的Helix格式一同作用）加以限制。

毫不奇怪，有些人因此建议利用DRM概念来确保和执行个人对其他种类信息的控制。纽约大学信息系统教授肯尼斯·劳顿提出建立一种技术与组织层面的基础设施，个人可以通过它将其个人信息的使用权“出售”给服务商，而基础设施本身将确保执行这些合约安排。网络法先驱劳伦斯·莱西格在其著作《代码》（*Code*）一书中认同了劳顿的理念，但更加明确地建议，应利用某种为知识产权开发的DRM基础设施来管理和执行信息隐私索赔。既然这些方法能应对个人信息控制权面临的挑战，它们也可能为数字化记忆提供一种合适的对策。有了这样一个系统，个人将能够在个人信息中添加元数据来详细限定可以使用该信息的人、用于什么目的、付出什么代价。嵌入在我们电脑操作系统中的规则将检查这些元数据并在确保某一信息的用途是可以允许的之后再处理该信息。只要全部或大部分技术设备服务商将这一系统的技术组件嵌入到他们的设备中，维权就得到了保障——甚至在法律上对信息隐私不予承认的某些司法机构那里也是如此。

DRM系统的反对者们常常将其丑化为消极否定的基础设施，以禁止和控制为中心。这是十分不公正的脸谱化评价，至少就考虑劳顿和莱西格的提议而言。从技术层面来看，限制未经授权的信息使用只是他们建

议系统的其中一个元素。至少与之同等重要的是一种技术和组织层面的基础设施，能够使那些希望利用信息的人便捷高效地与信息权限掌控方交易。其理念是建立各种信息市场，信息使用权能够在其中以远远低于现存市场价的价格交易，从而不仅促进了交易，同时还为权利所有者们进行价格体系试验提供了更强的动机。这将促成更加高效的信息市场，并因此总体上有利于社会。

然而，这一方法有一个连拥护者们也承认的重大问题。尽管该方法主要依赖技术机制来保障高效交易和有效维权，但任何此类DRM基础设施仍然需要一定的法律支持，尤其在限制某些尝试回避技术控制机制的行为时。举例来说，在知识产权法领域，包含在《数字千年版权法案》中的逆向工程禁令就扮演着这一角色。如果要利用数字版权管理来解决其他信息权利问题，那么逆向工程限制的范围就必须得到相应的扩展。

尽管利用DRM来保护和执行个人信息隐私权的方法遭到了猛烈而广泛的批评，但并非所有批评都是合理的。在针对DRM的为数众多，并且有时是意识形态层面的争论之中，有三类观点指出了数字版权管理在应对此类控制中存在的严重的理论和实践问题。

第一个问题集中于**DRM**系统的综合性。举例来说，在知识产权的使用场合中，信息内容可以通过加密而从其源头到终端用户一直得到技术保护。然而这样一来，内容就必须是可供人以视觉和听觉来欣赏的，因此就摆脱了DRM系统的保护限制。在这一“未加密时刻”，它是可以被录制的，并且可用于制造未经授权、未受保护的版本，供侵权者不受技术限制地大规模传播。当然，设备生产商一直不遗余力地试图将未加密的时刻变得越少越好。所以高清电视信号甚至在从数字机顶盒进入高清电视机的过程中也是加密的。但是，所有这些技术都还不能完全防止用户利用摄像机录制观看的电影或用录音机录制播放的CD音轨。当然，这样做会降低内容的质量。尽管性能卓越的录制设备能够制造清晰度惊人的副本，内容提供者们对此忧虑不已。然而，在个人信息的使用场合中，这个问题可能要轻微得多。人们在看或听那些受版权保护的内容时会导致“未加密时刻”出现，与之相反，个人信息常常是输入电脑供处理的信息，因而很大程度上会被限制在一种更易于保护的信息处理基础设施的技术范畴中。

第二个问题是，**DRM**的反对者们表示这样的系统绝不可能实现防破解。总有人能够突破用以保护信息的技术限制，甚至是在法律禁止此类行为的情况下。一旦这样的限制遭到破坏，信息便失去了保护。过去的事件似乎证实了这一评价。过去数年里，互联网上出现了各种能够轻

易获取的破解方法，能破解从DVD、PlayStation游戏到保护Adobe的PDF文件和苹果的iTunes歌曲的各种DRM系统。这一情形就像军备竞赛，更复杂的DRM系统最终都被更成熟的对手破解。大多数现存DRM系统的根本问题就在于，一旦被使用，它们就不可能针对破解做出调整，这使得信息内容极易受到侵害。设备通过全球网络相连或许能改变这一局面：只要包含某种DRM系统的设备通过因特网相连，它们就能获得更新，从而克服某些弱点，进而防止破解，至少原则上这是可行的。未来，这或许会使DRM系统更难被攻破。

第三个问题是，**DRM**系统的运作需要追踪用户的行为，并且当发现他们以未经授权的方式或为了未经授权的目的使用信息时加以阻止。这个关于利用DRM来控制个人信息的问题似乎更根本、更需要解决。姑且不论一个技术系统能否在多大程度上理解目的这样抽象的概念，任何有能力做出这种判断的系统都不得不观察用户是如何处理受保护信息的。本质上，它需要一种能够时刻监视用户并随时对未授权行为加以制止的技术系统。保护我们免受数字化记忆侵害的DRM系统本身就是通过无处不在的监视所构成的基础设施来实现，这是多么讽刺！用圆形监狱来保护我们免受圆形监狱般的社会侵害？或许未来十几年的技术创新能够消除其中的某些顾虑，但在那之前，全面而普遍的DRM系统仍将伴随着令人失望的先天缺陷。

第四个问题是，**DRM**系统管理的每一条信息可能的用途、花费和相关条件都需要详加定义。这样的元信息是该系统运作的前提条件。在知识产权范畴，产权所有者，常常是大型商业组织，十分热衷于雇用训练有素的员工来输入元信息并对其保持更新。在个人信息范畴，元数据必须由个人自己来输入。说服普通用户花几个小时来详细说明使用政策并且持续更新这一元信息，这似乎不是件容易的事。困难不仅在于输入使用政策对个人而言耗时费力，还有另一个难处：描述别人能如何使用一张图片或一首歌曲相对来说简单明了；而描述一条个人信息能否在什么条件下，为什么目的所使用则要困难得多，原因还不仅在于使用场合与目的的重要性。现在的DRM系统用特殊的语法和语言结构来描述某一特定信息能够怎样被使用，以怎样的代价使用。对于如何使用信息，目前的这些语法结构代表了一种相对原子论色彩（更不必说过于简单）的观点，忽略了场合和目的；这样的结构必须得到根本性改进，才能帮助个人高效便捷地把握住众多基于场合和目的的个人信息使用方式——这实在是个艰巨的任务。试想一个简单的例子：如果某人要对她社保号的适宜用途进行详细说明，她究竟需要多么具体？要知道，为了某一目的与他人分享社保号可能完全合理，但是换做另一情况时则有可能

令人忧心，即便是同一个信息处理者提出要求。

对于数字版权这一挑战有两种可能的解决方案。首先，我们可以将个人信息要求重新概念化，并将它们塑造成类似知识产权要求的形式。“产权化”个人信息的显著优势在于它能够更好地适应现有DRM系统中用以描述合法使用的语法和语言结构。但是正如我前面提到的，对个人信息的准产权要求几乎赋予了“所有者”对信息为所欲为的完全权力，消除了对信息重复利用的限制。

取舍之道

个人可以决定是否与他人分享信息，但对信息的使用场合和日后的用途却控制权有限。对于受版权保护的作品，我们已经接受了这种限制。如果我买了一本书，我可以在家或路上阅读，将其用于研究、在其中做笔记、卖给他人或撕成碎片。版权所有者几乎无法干预我如何使用这本书。对于个人信息，我们可能想用截然不同的方式来处理：我只会授权某一特定的信息处理者仅在某一特殊场合，为某一特殊目的使用我的社保号。标准的产权并不允许做出这样细致的区分。它们只是粗钝的工具，而它们的任务却似乎太过精细。另外，诚如信息隐私专家保罗·施瓦茨、马克·罗滕博格和知识产权领域权威帕梅拉·萨缪尔森等人精彩的论述，标准产权之所以并不可取还有其他原因，尽管在此我们认为缺乏适应性或许是最重要的原因。

如果使信息隐私要求适应技术性基础设施并不是个好主意，替代方案将是使技术与个人信息控制的理念相适应，并创造出更恰当的语言，以描述个人选择为某种目的，在特定条件下分享个人信息。这是个艰巨的任务，并且不是短期内能够解决的，但它可能是实现一种技术性解决方案，如一种DRM系统，以管理和执行个人信息控制的唯一途径。在那以前，这些全面的DRM系统对我们寻找一种应对数字化记忆的合适对策毫无帮助。

也许我们只是把DRM系统想象成面面俱到了。我们或许能利用现有科技能够实现某种系统，它既能提供一个涉及面相对较小的解决方案，同时又能保留效率。

普林斯顿大学信息安全专家爱德华·费尔顿及其同事们正是提出了这样的建议：一个有助于双方协商个人信息使用和执行控制的“隐私管理系统”。与纯产权系统不同，费尔顿的这一系统建立在当事各方直接的共同许可上。有趣的是，这一系统还允许当事各方对使用政策保密，从而使各自免于外部压力。它基于透明性，这是传统DRM系统所不具备的一个特点，极具魅力又令人好奇。然而，这一系统的核心魅力在于它不需要一个权利管理的中央基础设施，并且因此免于建立监视系统。

这些有意思的观点对设计过度的方案而言是有益的解药。也许，有限的“隐私管理系统”在考察现有的信息隐私行为和实践方面，为我们实现一个切实可行的技术性解决方案提供了一条更具潜力的长期途径，这一方案将保障个人信息得到控制。

总之，不论是全面基础设施还是更有限的系统，技术性解决方案较之前文提及的两种可能的对策——数字化节制和信息隐私权，的确具有某些优势，但仍有缺陷。我们该何去何从？尽管我们每个人都有个人偏好，但这些对策中的赢家并不明朗。每种方法都有其优点和缺点。

数字化节制能带来行为变革，而无需向个人施加外部约束——无论是法律还是技术的，但它是否能成功应对数字化记忆的挑战还是未知数。信息隐私权代表着一种更有组织性、更健全的对策，并且已经在世界上众多司法机构得到了成功应用，但是其效果仍然存疑。管理个人信息控制权的技术系统看似在解决执行（因此也是效果）问题，但技术系统的运作规则、哪些信息使用行为是合法或非法的，这些问题对专家而言尚且复杂得惊人，对普通的个人而言更有可能无从下手，即便这些系统的其他问题能够得到解决。这些方法中的每一个都代表着一种旗帜鲜明的行为规范机制。数字化节制利用行为规范（以及一点市场手段），信息隐私权依赖法律，而权力管理系统主要采用技术架构。由此看来，它们似乎代表着在我们的社会中已经存在的行为矫正范式方法。

然而，我对它们的兴趣有限。它们能不能恰当而有效地应对我们所面临的数字化记忆的挑战？

重要的是，它们的核心思想在根本上都是关系性的。它们都预先假定我们面临的核心挑战是信息从发送者流向接收者，从一个人到另一个人，或从一个组织到另一个组织。它们假设发送者决定着接收者是否应该接收信息，且应该在什么条件下使用该信息。每一种对策的目标都是促进或塑造信息分享的关系条件。这一点也不奇怪。这三种方法都植根于信息隐私的观念，其核心就是一个关于约束信息分享的关系概念。

取舍之道

就像我在前面解释的，遗忘的消亡带来了两方面的挑战。首先是对相关信息权的挑战或失去信息控制权的危险。我描述的这些性质相似的方法或许可以应对这一挑战。至于它们是否也能应对时间和空间上“圆形监狱”这样更具社会性的危险，我就不那么确定了。另一个问题，我称之为数字化记忆的时间挑战：即太多的数字化记忆可能会压垮人类的思维能力，决策能力、应变能力和怀旧能力——这种威胁并不是关系性威胁。

我们的数字历史会阻碍当下的能力，这一威胁要比他人拥有我们的

信息大得多。我们即使没有与他人分享信息也会深陷其中，就像珍妮阅读自己与约翰的旧电子邮件时那样。我担心的是，数字化节制、信息隐私权和隐私DRM所代表的关系性方法并不适合用来应对这一挑战。这不是它们的设计初衷。所以，无论这些方法在保护信息隐私方面多么有用，它们并不能直接应对全面数字化记忆带来的“时间”挑战。

是否存在这样一套替代方案，它们不那么明确专一地建立在关系概念之上，因此能够更好地将数字化记忆视作一个信息问题（而非关系问题）来解决？如果我们同意数字化记忆就是应该解决的问题，即一系列存储于数字媒介并因此封存在时间中的信息，那么，一套包含三种方法且更有针对性的对策——建立在社会规范、法律和架构之上，便呼之欲出。

这一套对策相比前面提到的那套更为熟悉的对策思辨色彩更浓。但是由于较少关注信息的关系维度，相信它们或许能更好地应对数字化记忆带来的“时间”挑战。

对策4：调整人类的现有认知

当我最初那篇关于数字时代遗忘消亡的论文在Slashdot网站上被讨论时，一位欧洲人工智能研究者朱利安·特盖里亚斯对我的“诊断”表示了认同，但提出了截然不同的“治疗方案”。他写道：“与其那样，我们不如调整自己的文化，以适应现代科技的存在。我们只需假设，人们可以做出改变并限制自己只关注最近的行为和意见。”特盖里亚斯的意见得到了哈佛大学伯克曼中心的同行达娜·博伊德的呼应，她确信“人们，尤其是年轻人早晚会想出应对机制。这将是一种转变，不是政府或某种技术实体干预的产物”。特盖里亚斯和博伊德以及相当一部分学者，并不否认我们正在向一个不再遗忘的世界迅速靠近，但他们相信人类会调整自己的认知过程以应对数字化记忆。这是一个让人兴奋的观点：如果我们能有意识地忽略过去，接受人类是永恒变化、绝非一成不变的事实，那么所有信息的长期存在将不再具有那么大的威胁。这样的认知调整将消除决策不明或反应不及时的潜在危险。

更确切地说，认知调整之所以这么吸引人，原因至少有两点。第一，它不需要通过采用新法律来改变社会，或发展和使用新的技术架构。相反，必要的改变发生在我们的头脑中。第二，通过重塑我们思考、评估和决策的方式，人类得以用适应能力来应对数字化记忆，而这正是生物演化的根本要素。这将使我们再度与变化——自然界最基本的概念相同步。我们可以借助零假设来处理这一问题，即不去尝试通过社会手段加以解决。最终，我们会调整自身的认知机制，使自己适应数字

化记忆的世界。记忆仍将伴随着我们，但我们经过调整的头脑就像拥有特氟龙涂层的锅一样，它会在记忆令我们困惑之前将其忽视。

特盖里亚斯和博伊德十分明智，所以没有想当然地认为这样的调整能够轻易实现。我确信，他们肯定预料到我们在达到目的之前将经历一段崎岖不平的过渡，从中获得必要的认知更新。我对此毫无异议，但真正的问题是这个过程究竟要多久。与那些认为认知调整会很迅速的人不同，我对这一变化能否在一两代人的时间内发生持怀疑态度。一些心理学界领军人物的观点更助长了我的怀疑。

大卫·沙克特认为，人类管理记忆和遗忘的方式不能轻易和随意改变，而是深深植根于我们大脑的内部运作之中。另外，J.R.安德森认为，人类处理时间（尤其是通过遗忘）的方式产生于长期以来对生存环境非凡而有效的适应。

如果人类能够将记忆任意放进某种时间观念中，那么那些被长期记忆困扰（如只能记起创伤）的人就能轻易地治愈了。哎，可惜我们没有洞悉实现这一简单解决方案的秘诀。不排除这一点会得到改变，但从我们今天知道的情况来看，我们似乎不大可能在数字化记忆的阴影下迅速地实现进化性适应并发动一场相对仓促的认知革命。

然而，如果我们实际上可以为适应数字化记忆而对大脑做出调整，情况会怎样？这样的认知改变仍然没有解决数字化记忆的另一个根本缺陷：不完整性。即便是在数字时代，我们所有的交流也并非都是数字形式的——更不用说我们的思绪和做出决定前的评估与权衡利弊了。这样的不完整性在我们日后重温某个决定时会带来问题。浏览数字化记忆不会为我们提供全面的背景和完整的信息，而我们需要这些来重建，并因此理解过去所做出的决定。然而，没有了这些信息，过去仍将困扰我们，除非我们也在认知上适应了这一挑战，或者干脆完全对过去弃之不顾。可是，一旦我们开始权衡和评估应该使用哪些，应该抛弃哪些，我们就不得不与数字化记忆中包含的偏见作斗争（同时也受其影响）——能够数字化的信息得到了保留，而一切无法被数字化的则被忽视。我们似乎不大可能仅仅依靠认知训练来完成这个任务。

这种观念的拥护者们可能会争辩，就像启蒙运动和现代性教会了我们批判思维一样，几个世纪以后的人类肯定会嘲笑我们今天所面临的数字化记忆难题，因为那时他们已经学会了如何应对。这可能是对的，但是我对几百年后的事情没那么大兴趣。我仍然对人类如何应对这两者间那数十年痛苦的（或许还是粗暴的）改变深感忧虑。诚然，我的怀疑不过是揣测，而另一种可能（即人类迅速调整了认知能力）尽管显然不大可能，但也不是完全不可能。但是，除非我们揭示出认知调整为何行之

有效的其他原因——在自我控制和自我纠正的个人层面解决数字化记忆的信息问题，寻求应对数字化记忆的替代对策或许是更富有成效的策略。

取舍之道

在拥护者的核心思想里，认知调整承认全面数字化记忆，却致力于限制数字化记忆对我们的决策带来的影响。如果说这一目标太好高骛远，因为人类只能考虑易于处理的信息，那么法律形式的社会规范或许能限制数字化记忆的内容，进而削减包围着我们的海量信息。这将我们引向了信息生态的理念。

对策5：打造良性的信息生态

信息生态是一种有意的约束，规定了什么信息能够被收集、存储，并且能被谁记忆、记忆多久。回想一下荷兰人口登记的悲剧，其可怕的结果提醒了我们，对政府在人口登记中的行为加以限制并不能防止不确定的未来；但是，一开始就不收集和存储信息却能做到。如果没有人口登记中关于信仰和种族的信息，纳粹也不可能对其进行如此可憎的滥用。对数字化记忆进行限制不仅能确保其日后免于外来入侵的危害，还可以在未来当我们的社会受到蛊惑，或者出现一些人利用数字化记忆的信息宝库进行不正当记录、歧视和威胁不认同主流价值偏好的人时，保护我们不受伤害。我们都见证了“9·11”之后，即便民主制非常健全和强大的国家也忍不住怀着加强国家安全的美好愿望（常常是虚假的）搜集和窃取数字化记忆。正如荷兰案例所启示的，收集个人信息天然就带有风险，因为我们不知道它们在未来会被用于什么目的。信息生态规范是应对不正当信息使用（或重复使用）的一剂良药。

信息生态规范详细规定了信息能在数字化记忆中保存多久，这并不是什么新鲜东西。人们经常看到它们以信息删除令的形式出现在各种不同领域和场合。

一旦满足了某些条件（包括守法表现和随着时间流逝），犯罪记录可能会被消除——封存或销毁（虽然后来许多司法机构倾向于封存而非删除信息）。类似地，经过一段时间后，如果不再需要某些从目击者或嫌疑人身上采集的DNA信息，该信息就会从DNA数据库删除，至少在美国某些州情况如此。同样的情况发生在最近一次罕见的决议中，欧洲人权法院在这份决议中规定，无辜者的DNA信息不得永久保留，迫使英国修改了法律并强制执行信息生态。

除了这些关于领域和场合的具体规则对信息删除和消除进行着硬性规定，信息生态的大体原则已经被包括在许多信息隐私法规中。例如，

《欧盟隐私保护指令》第6条规定各成员国法律必须确保个人信息“以一种允许对数据主体进行识别的方式保存，期限不得超过收集或进一步处理数据之目的所需的必要时间”。这样的信息生态规范不仅存在于正式法律中，还存在于各个信息处理技术开发公司的隐私指导原则中。其中最大的微软已经将信息生态突出地包括进其“产品与服务开发隐私指导原则”中。在其点明主旨的标题“缩短数据保留期限”下，他们声明：“用户数据的保留期限应以达成业务目的或满足法律要求所需的最短时间为准。”

这些规范带来的影响一目了然：一旦个人信息达到了其使用目的就应当被删除。例如，一旦相应的账单得到了结算，信用卡公司和电信服务商就必须从其数字化记忆中清理掉个人信用卡交易或个人通话记录（或至少抹掉个人身份标识，如姓名或此类记录的帐号）。某种程度上，这样的强制性信息生态是目的限制原则的必然结果：如果因某种目的将个人信息委托给某人，那么一旦该目的达成，对方继续使用该信息就是非法的，必须将其删除。

在学术文献中，已经有人为更加严格、全面的信息生态授权提出了建议。耶鲁大学法学教授杰克·巴尔金呼吁将其所谓的政府失忆制度化，即“每隔一段时间对某些种类的数据定期进行销毁，除非有充分的理由对其予以保留。”几年前，我的同事和我主张建立一套全面信息生态法，它将作为“信息减速器”，迫使信息处理者们放慢信息收集和存储的节奏。

无论是作为针对某一领域的法律还是通用信息隐私法的一部分，这些规范都有效地维护了遗忘。由于它们都是针对信息存储者的立法委任权而非要求个人为维权诉讼，所以信息生态规范为应对数字化记忆的挑战提供了一种更直接有效的对策。不幸的是，这种策略有两大缺陷。一个是概念性缺陷，一个是政治性缺陷。

概念性缺陷在于强制性信息生态是个相对粗糙而二元化的工具。法律法规明文规定了什么信息应该被删除以及何时删除，而没有涉及个人决定的后果，就像信息隐私权那样。这就需要建立一个一般性，而不是灵活可变的规范，因为法规必须精确到没有多少可供解读的余地。毕竟，人们必须先理解规范才能遵守规范。于是，大多数现有的信息生态规范，对一些具体的疑难情况进行了限制，比如DNA信息：而对大多数其他同样可能影响我们做出理性决定的数字化记忆事件网开一面；或者它们和某种关系意义上的权利概念相关联，如《欧洲隐私保护指令》第6条。前者限制了信息生态规范的潜在范围，而后者正如我所述，未能有效地应对数字化记忆带来的“时间”挑战。

政治性缺陷是指，尽管数字化记忆正迅速崛起，信息生态规范近年来却很难获得立法。相反，在“9·11”过后和随后的“反恐战争”期间，保留信息而非删除信息这种反其道而行之的法规却层出不穷。银行必须对进行国际转账的用户收集和存储个人信息，表面上是协助政府禁止对国际恐怖主义进行经济支持的活动。所有向美国输送旅客的航空公司都必须在飞机进入美国空域之前向十几个美国执法机构提供每名乘客的详细个人信息，包括电子形式的地址和信用卡信息，以便这些机构将乘客数据与潜在嫌疑人数据库进行对比检查（甚至连参议员特德·肯尼迪也曾经被收录在这样一个政府“禁飞”数据库中）。类似地，美国和欧洲的电信服务商现在也被迫将其用户电信连接中的交易数据保持数月，甚至数年之久，仅仅是考虑到执法机构或许会查阅。最近，为了进入欧洲的一个商业飞行模拟器，我不得不提供详细的个人信息。政府现在也要求对这样的记录进行收集和存储。

信息保留法的兴起是由一系列不同的因素造成的。首先，后“9·11”时代安全之忧的表面源头——全球恐怖主义，是很难直接用矛头对准的。恐怖分子分布广泛，对其组织的其余部分知之甚少，并且常常相对独立地行动。尽管过去几十年也曾出现过某些剧烈的先例，执法机构还是将全球恐怖主义视作一种新威胁，并认为以现有的手段和工具无法与之斗争。这导致执法重心从解决犯罪转向了预防犯罪。因此，执法机构已经意识到全球信息和通信网络中流动的信息量之大及其潜在价值。有一种观念认为，只要正确解读和关联信息，就能在下一次袭击发生前成功侦察到恐怖活动。由于执法机构不收集相关信息，并且常常缺乏相应的信息保存能力，所以新信息保留法被视为解决这一局面的良方，即通过强迫私人公司保存顾客信息来便于执法机构将来查阅。

无论这一论点正确与否，近年来信息保留法的复兴不仅逆转了信息生态规范数十年来在遗忘方面取得的进步，还助长了一种反对信息生态的政治环境。如果政府为了公共安全而认同对个人信息进行广泛和长期的保留（就像小布什总统任内所做的），它就不大可能积极地在私人领域执行现有的信息生态规范。

信息生态规范之所以会在困难时期没落，还有另一个同样重要的原因。针对近来一连串私人领域的腐败案件，美国国会用透明度法规以及其他措施进行了回应，这与过去数十年内最臭名昭著的公共领域丑闻水门事件在发生后采用相同的应对机制。无论是笨拙的《萨班斯·奥克斯利法案》，过时的《信息自由法》抑或食品包装必须印刷营养信息的要求，透明度似乎成了我们确保良好管理的首选措施，不论在公共还是私人领域都是如此。这一趋势不仅在美国显而易见，并且在许多其他社会

也存在。例如，欧盟的环境法就包含重要的透明度与自我报告成分，工作场所的安全规范中也是如此。甚至在传统上透明度立法稍显欠缺的国家，比如德国，政策制定者也于最近对透明度法规有了兴趣，建议从经理薪水到政客的董事会资格等事无巨细都采用透明度法规。尽管某些学界人士已经对透明度法规的过度使用表示警惕，透明度仍然被政策制定者和公众视作减少腐败和强化监督的首选机制。

总的来看，信息保留法的复兴和透明度法规的涌现反映并促成了一种反对信息生态规范的氛围。但这并不表示个人立法者们没有提出信息生态的措施。比如，2008年，马萨诸塞州众议员比尔·斯特劳斯提出了一个法案（H4822），该法案强制公司在24个月后删除行为营销信息（基于消费者偏好——即何时会购买什么）。然而，主流的政治气候没能允许这样的信息生态规范获得立法。

取舍之道

目前为止，我们已经审视了两种明显不同的对策。一种限制数字化记忆（信息生态）的信息量，另一种则不干预数字化记忆，而是提倡从认知层面将其中相当一部分忽略。每种对策都有其价值，但都没有带来突破。认知调整似乎对人类来说太难实现了，至少就短期而言；信息生态规范既面临着敌对的政治气候，又面临着内在灵活性的缺乏。两者都包含一个要求，人类应该对他们做出决策所需的信息加以限制。或许这一假设存在漏洞。除此之外的第三种替代对策，完全语境化，则提供了一个截然不同的视角。

对策6：完全语境化

如果说数字化记忆包含的信息不是太多，而是太少呢？目前为止，我已经指出，随着越来越多信息被添加进数字化记忆，数字化记忆将会因为向我们提供大量或许应该被遗忘的信息而给我们的决策带来困惑。于是，我们逐一审视了应对数字化记忆威胁的可能对策，有些专注于减少信息的人际流动，即前三种对策都涉及的关系维度；有些专注于减少数字化记忆的信息量，即后两种对策核心关注的信息维度。

然而，正如认知调整的拥护者所反驳的，人类不是完全有能力评估不同时期互相矛盾的信息吗？这不正是我们一直都在做的吗？如果我们能够避免大多数数字化记忆收集信息时固有的偏向性，难道认知调整还不能行之有效吗？如果把包括某一历史事件的背景在内的所有信息都保存在数字化记忆中，而不仅仅是相对较小的（就现状而言）子集，难道我们还不能适应吗？数字化记忆之所以成为问题难道不正是由于其有选择性地记录历史信息，因而无法令用户总揽全局吗？

这些完全语境化的论点不仅是认知调整的变体，而且是对布林为呼吁建立更透明社会所作论述的扩展。在他所说的这种社会中，所有人都能获取有关任何事的全面信息（至少原则上如此），个人信息控制被普遍透明度替代。

取舍之道

为了实现完全语境化，我们需要建立一种技术基础设施，它能比现有的技术更全面地收集、保存和提取我们生活中的各种信息。我们必须放弃模拟存储和未经录音的对话，并且不得不成为戈登·贝尔那样的生活记录者，将所有经历和最终做出决策所需的信息都记录下来。但是，我们也必须超越戈登·贝尔，并利用各种工具与他人分享信息，以获得真正的、平等的透明。简单地说，我们需要实现真正数字化记忆的技术手段。

信息隐私专家们或许会辩称，完全语境化是在创造一个邪恶的监视设施。但是奥斯卡·甘迪提醒我们，监视是一种不对称的观看权，一种必然会给予监视者相对于被监视者优势的权力。而布林认为，在一个所有信息都透明的世界，监视便失去了这种权力。同时，如果数字化记忆的基础设施能够以一种权力分散的方式组织起来，酷似戈登·贝尔使用的那些设备，或许就没有哪种单一实体能够对其全盘控制，无论它的权力多么强大。虽然多少有些挖苦意味，但生活记录者们正确地指出，他们倡导的不是监视（一人观看他人），而是反监视（我们观看自己）。

然而，正如我在第4章中提到的，这一论点存在严重的缺陷。完全透明不一定有益。布林自己便指出，这样的完全透明会夺走我们的匿名权，令每个人都感觉自己生活在一个小村庄，事事被人看在眼里，所以不得不谨言慎行。“在这个村子里，你不敢对邻居冷言冷语的根本原因不是害怕受到惩罚，而是因为你知道肯定会有人向你母亲告状，从而给家族蒙羞。将来，当任何公民都能进入那时的万能数据库时，我们的‘村子’将会包含数百万人，而且陌生人的母亲和你不过只有一封电子邮件的距离”（原文强调）。布林将某人的父母视为有权者（相对而言），这一提法把他善意的隐喻重新带回到边沁的圆形监狱概念中那种社会控制机制，一种压抑的监视架构。正是因为这个原因，所以很多人不愿与他人分享信息。甚至连生活记录者戈登·贝尔也坚定不移地将其信息宝库留给自己。如果所有人都能获取所有信息，那么有些人极有可能会拥有各种技术、组织和经济手段来更好、更深入地整合信息，从而导致一个新的有权者与无权者格局出现，与布林的理想主义憧憬相背离。

即使是最好的技术系统所能记忆的信息量，也面临着一些理论上存

在的限制。如果我们相信这样一种系统中的信息是完整的，那就是自欺欺人。首先，以我们现有的能力还远远不能将身边的信息都进行数字化记录。我们必须在技术传播和认同方面实现量子跃迁般的飞跃才能使数字化记忆的全面性达到一个合理的高度。同时，即使所有通信和外部信息都得到了记录，数字化记忆仍将缺少一个无法被数字化的（至少目前如此）重要信息源：我们的思维。如果只记录外部信息，而没有记录我们内在的思想，数字化记忆仍将在根本上是不完整的，更不用说熵定律对我们信息存储量的限制了。只要数字化记忆仍然存在明显的不完整，它就不仅不比现在的情况好到哪里，比如人类的遗忘行为，甚至还要糟，因为关于存储什么信息，舍弃什么信息的筛选过程不是基于人脑的内部运作，或纯粹随机，而是受到技术工具记忆能力的限制。

即使技术上可行，完全语境化仍需要我们持续关注。为了完全理解一个过去做出的决定，我们不得不潜入一片相关数字信息的汪洋大海，其中还包含着它们各自的语境。因此，由于人类注意力的广度有限，完全语境化可能无法使我们全面地理解过去。

最后，完全语境化还无法处理第二维度的问题，即及时决策和反应的问题。作为人类，我们是线性时间的俘虏。即便是在一个精确的思想实验中，我们也不可能回到自己的过去，原因是我们不能否认这样的事实，即我们人类一刻不停地在学习、演化和改变着。即使我们面对的是一份对过去经历的精确记录，我们也不可能屏蔽掉那些后来已经在脑海中发生的改变：获得（或失去）知识、价值观改变、个人偏好变化和情感经历。我们在阅读一年前度过的书时就能体验到这一点：我们会有新发现，对同样的事物会有不同的解读。同样，即便在完全语境化和时间无限的条件下回顾曾经的某个决定，我们也不可能重溯并理解当时做出决定的过程和原因。相反，我们可能会不认同自己当初的决定，问自己为什么会错得这么明显，做得这么不对。简而言之，即使完全语境化能重现信息的语境，它也不可能带我们回到过去。

这些缺点如果集中到一起，将会显著地降低完全语境化可能具有的价值，使得它在应对数字化记忆方面的魅力大减。

目前我已经提供了六种应对数字化记忆的对策。由三种对策构成的第一套对策专注于个人控制信息分享的能力，反映了数字化记忆的“信息权力”维度。第二套对策的核心关注是人类利用信息进行决策的过程，与“时间”维度相呼应。这些对策各自利用了一种独特的机制来规范人类行为：社会规范和个人自制、正式法律，以及（技术）架构（见表5—1）。

表5—1 应对数字化记忆的可能对策

	信息权力（包括信息隐私）	认知、决策和时间
个人	数字化节制	认知调整
法律	隐私权	信息生态
技术	隐私DRM	完全语境化

尽管每种对策都看似处理了某一方面的问题，但这六种对策中没有一个能够有效地反制数字化记忆的潜在威胁，或许这暗示着我们面临的挑战是多么复杂。表5—1左列的方案（主要处理数字化记忆的“信息权力”维度）似乎更成熟完善，这一点也不奇怪。它们来自于数十年的研究、行动、辩论以及针对已知信息隐私损失的立法。右列方案（处理“时间”维度）相比之下的缺陷也并不惊人。只是由于近年来数字化记忆的兴起和遗忘相应的消失，这一挑战才变得突出。

总体来看，我们或许可以得出两个结论。第一，为了应对数字化记忆的挑战，我们也许应该放弃对某种完美答案的寻求，转而务实地致力于寻找某种混合方案以确保这一挑战的所有必要方面都得到解决。第二，设想出更多的创新性方法或许是表格右列最需要的，因而也是解决数字化记忆的“时间”维度问题的需要。这样的一种方法正是下一章的焦点。

---

[1] 即通过对信息进行强有力的组合与再组合来产生创造性成果，第3章也有提及。——作者注

06 给信息一个存储期限：应对数字化记忆与信息安全的关键对策

存储期限并不是强制性的遗忘，不是让我们被迫去选择，而是通过存储期限让我们能对信息的寿命做出应对。它将成为我们日常生活的一部分，让我们深刻意识到一个人类已经无意识地默认了上千年的道理：数量不等于质量，“好”信息不等于“滥”信息。

存储期限，不是二元化的选项——有用或无用，而是提供一种可以保留一段时间而非永久保存的选项。这种机制能够在未来的某个日期删除信息。由于提供了记忆和遗忘之外的选项，这种机制显得更有效，并且更接近人类的遗忘能力。

信息存储期限的核心，不是借助技术手段将数字化记忆的问题推离我们的意识，恰恰相反，它让人们觉察到遗忘的价值和重要性。

存储期限逐渐为我们担负起了理解数字记忆的重担，为人类和社会保留了“快速反应”的能力。

cookie的警告

如果想要讨论数字化记忆所带来的挑战，我们就不得不提到如今这个重要的现象——遗忘的消亡：过去曾随时间而忘却的，如今能被我们准确地记起。在我们与彼此以及与技术工具之间互动时，用数字化记忆保留信息已经成了默认选项。在今天，记忆是如此的普遍和广泛，不再需要有意识地努力，不再需要多花哪怕一点点的时间、精力或金钱，我们就能在数字化记忆中保存信息。

我认为，我们可以重新建立一种平衡，使遗忘仅仅变得更容易一点点——仅仅将这种默认值调整回千年前的状态，从永远牢记到逐渐遗忘。我们可以用多种方式达到这一目标。我们不妨来看看这种可能的途径吧：将提升人与社会的觉察度、技术工具与法律支持共同结合起来。它是如何区别于现存的信息权力和控制手段的呢？可以想想现存的cookie。每当我们浏览网页，网站就会发送用来检测我们在做什么的小型数据包。

十年前，我写了一篇关于cookie的文章。在这篇文章中，我解释了网页浏览器，如火狐或IE，可以被设计为询问我们是否接受或拒绝随输入而来的cookie。但是，浏览器却不能告诉我是否已把信息发送到其他网站，这让我感到惋惜。我在文章中指出，这种发送在暗中发生，并且明显违背了已经广为接受的透明原则以及信息隐私权的主旨。在意识到这种权利的存在之后，我立刻警觉起来，因为我的浏览器可以在询问我之前就把我的个人信息分享给其他网站了。

鉴于数字化记忆，我现在意识到当时的推理是不完整的。就像在表5—1中提到的三点，我的论证集中于用户与提供信息的信息处理者的连接。它根植于信息共享的关系性概念之中，一个我们许可和赞同的世界。它所缺乏的是一种对cookie警告的欣赏，弹出的窗口让用户拒绝一个收到的cookie，而不是让他们拒绝是否与别人分享cookies。在本质上，如此的cookie警告确保一些信息不会进入我们的数字化记忆，以此来帮助我们限制整体的记忆。并且，这些收到cookies的警告会使一些用户烦躁，但也会产生积极影响：使用户对数字化记忆永无止境的真相引起持续的警觉。与DRM之类技术框架不同的是，cookie警告不能为用户解决问题，但让用户面对问题。因此，cookie警告依赖高频的对用户的骚扰作为提醒手段。

最终的结论是，**cookie**警告对于限制数字化记忆是有贡献的，它的弱点极度的二元化。用户的选择是一种“前期投入”：当接收到一个cookie，用户可以权衡利弊并决定是否将其加入数字化记忆。但如果用户想要仅仅保留cookie一段时间怎么办？更现实的问题是，如何避免二元化的选项（不是有用就是无用），进而提供一种可以保留一段时间的选项而非永远保存？这种机制能够在未来的某个日期删除信息（cookie警告并没有提供这一功能），会因为提供了记忆和遗忘之外的选项而更有效，并且更接近人类的遗忘能力。

信息的存储期限

我将提出一个重要的概念，信息的存储期限。在数字领域内模仿人类遗忘的可能方法之一是，把我们保存在数字化记忆中的信息和一个存储期限相关联，让数字存储设备可以自动删除那些达到或超过存储期限的信息。我相信，这种极简单的形式足以再现真实记忆的运作方式，尽管我们还能想到很多其他方法。无论是简单还是复杂的方法，都拥有一个共同的核心元素：提醒我们，并让我们面对信息在时间上的有限性。换句话说，信息必定与某个时间点或某个时间段紧密相关，并且随着时间的推移，大多数信息失去了他们作为信息的价值，比如隔夜的报纸，或者是老掉牙的笑话。

当用户存储了一个新建文档时，除了选择文件名和它在硬盘上的地址以外，他们会不得不选择一个存储期限。如果没有选择存储期限，用户将不能保存文件，就像他们不能在没有起文件名的情况下保存文件一样。基于这一前提，用户的电脑就可以自动做剩下的事：管理存储期限并清除过期的文件，可能每天一次。当然，用户可以利用软件上的灵活性去改变存储期限，以防信息在还没有失去价值之前就被删掉了，或者

因为某些信息变得更重要，而超越了它之前的预期寿命。偏执的人甚至可以拥有一种能够在信息接近过期时警告他们的工具，因此他们在十分迫切的情况下可以决定调整这一时间。但是，通过使我们的电脑删除那些到达存储期限的文件（就像我们清理过期的食物），日常生活又重新引入了遗忘的机制，并将我们的默认选项转变为曾经的那种在人为控制之下遗忘模式。

使用存储期限的成败将取决于用户的体验，取决于用户是否能够轻松界定存储期限。对存储期限的需求应该会促进，至少是在一段时间内促进用户去思考信息的存储寿命，但它不应该用复杂和笨拙的用户界面滋扰用户。允许用户使用模糊的时间将使这一过程变得更容易（比如说一个月或今天或一年），也可以是一种仔细思考和选择过的预设或默认，也许根据文件类型、存储位置（有些文件夹放置更重要的内容），或者对文件内容的（有限的）语义分析。

存储期限并不是强制性的遗忘。他们的目的是为了唤起人类的觉察和行动，要求用户哪怕只是在有限的时刻做出应对。也许对其他的反应相比（比如敏感词或信息隐私权），存储期限并没有基于一种过度理想化观点，如高层次用户会放弃网络或诉求信息隐私权。它也没有基于一种同样理想化的纯技术方案观点。与这些不同，受我们支配的强大技术工具只是被用来提醒我们信息的价值不是永久的。更重要的是，我们不会被迫去选择，而是因此能够对信息寿命做出判断，它将成为我们日常生活的一部分。我们可能也会意识到一个人类已经无意识地默认了上千年的道理：数量不等于质量，好信息不是滥信息。

元信息，设定信息的寿命

从技术层面讲，存储期限是相对容易实现的，它只是另一种和信息相关连的元信息。我们的技术设备已经能够存储更海量的元信息。文件名、创建时间、修改时间只是其中的三个例子。数码相机会存储每张照片的元信息，从曝光时间、光圈到胶片速度。大多数照相机在照片被拍摄时增加了日期和时间。很快，我们就会看到地理的元信息被加载到我们拥有GPS功能的照相机或电话上。我们的音乐软件能够可以为我们的音轨管理元信息，从名称、乐队、类型到歌词、唱片集、播放列表，到我们的关键词和喜好。是的，DRM系统也已经嵌入到了我们的大多数媒体播放器之中，提供（或被迫提供）了作为播放权限和版权的元信息。

元信息不是限制我们计算机无关紧要的小玩意。它们中的许多推动了Web　2.0的形成：在互联网上我们为喜欢故事点赞，为帖子评分，推

荐好卖家或者评论维基百科上的信息。这些都是各种各样连接着网上其他信息的元信息。存储期限可以是在数字化记忆中增加的另一种元信息——关于信息寿命的信息。它可以“粘”在它指涉的信息上，保证如果一个人复制了一个文件，期限信息会随着它一起被复制走。在这种背景下，一个小的应用软件可以定期清理存储期限。在我们的许多数字设备中，拥有这种功能的框架已经存在。例如，我们的个人电脑每隔几天或几周自动清理我们硬盘上的信息和辅助文件。清除过期的文件可以只是一个附加的日常任务。

取舍之道

为了保证这种技术框架的实现（因此人为设定存储期限是必须的），法律的规范必不可少。强制设备包含支持存储期限的代码，要求用户在保存时设定存储期限，并自动清除过期的信息。这不会是第一个规定软件功能的法令。现存的法规已经通过强制软件制造商遵从特定的安全和可靠性原则（从数字签名和飞行控制系统的投票机制），或甚至要求软件遵从某些人体工程学原则。我在这里所提出的存储期限只是再增加一项要求，而且是一个相对容易实施的要求。

9个月，不断缩短的存储期限

存储期限限制了公司甚至政府获取用户和公民信息的数量，这会使他们感到麻烦。然而他们仍然会热烈欢迎存储期限。今天，商务和政府信息处理者保留了大量的数据仓库，但是缺乏那些数据仓库信息是否仍然有价值的线索。如果没有这样的线索，Google或亚马逊的处理器会采用复杂的算法去模糊推断用户在搜索或为用户推荐书籍时系统应该动用哪些信息。这些猜测工作天生的不准确性常导致错误的结果（对书籍的错误推荐或网页检索结果偏离用户的期望）；对政府执法者而言，则会导致误抓好人。如果在合适的时机使用存储期限，被处理器保留的将是被充分优化之后了的数据库。如果大多数用户设置了存储期限，这种信息质量的增加额外成本，即使有的话也是非常少的。

用户也会信任那些不会滥用个人信息的网络系统服务。如果网络服务商辜负了这种信任，用户可以选择其他服务商。这一法则对互联网巨头们也是适用的。哥伦比亚大学法学教授蒂姆·吴（Tim Wu），评论道：“一旦你失去了对Google的信任，这将是他们的末日。”对于竞争者也是同样，可以利用信用丑闻宣传它们的相对优势。

当然变换服务商需要成本，成本大小取决于其他因素。改变e-mail账号是很困难的——你必须通知朋友或合作伙伴，可能会导致名片作废，甚至不得不去熟悉新的e-mail界面。相比之下，更换搜索引擎就容易得

多。

网络服务商例如Google、微软、或Yahoo！对上文提到的这些因素极其重视。在2007年，Google宣布它会开始对搜索请求做匿名处理。它这样做至少争取了用户的信任，但就在几个月前，Google就险些辜负了这一信任。在2005年，美国司法部请求调查数十万的顶级网络搜索公司的搜索请求。当时大多数搜索公司毫无异议地同意了，而Google却在联邦法院对抗这一要求并最终取得胜利。但它恐怕不能抵抗后续的审查。这使搜索巨头陷入了困境。它已经在数字化记忆的观念周围建立起了堡垒，并雪藏了它获得的几乎每个比特的信息。这使它成为了想索取大量用户信息的政府机构的主要目标。全世界的用户已经开始对Google的强大信息权力感到不适。任何与美国政府大量共享用户信息的行为都会使Google对用户的承诺岌岌可危。Google的管理层没有更多的选择，只能与用户一起对抗政府。但这只是赢了一场战役，而非整个战争。为了在政府能够成功取得这些信息的情况下长期守住信用，Google已经决定广泛为搜索请求使用存储期限：现在其中的个人标识在最多24个月后被自动删除。

取舍之道

Google的声明，给予了竞争对手一个用更短的存储期限来挑战互联网巨头的机会。在2007年夏，微软和搜索网站如Ask.com联合声明，它们将在18个月后丢弃那些来自用户的查询请求，比Google缩短了25%。就在大约同一时间，Yahoo！宣布它的丢弃时间为13个月，比Google快一倍。在那一年以后，Ask.com甚至在网页上使用了一种存储期限机制。仅要鼠目标一下点击，用户就可以如愿地删除他们的搜索历史。在2008年7月，新的搜索引擎Cuil宣布它根本就不会存储个人查询请求信息。在2008年9月，Google以把时间调整为9个月做为对抗。

这种动态的竞争为信息的存储期限上了激动人心的一课：就如同用户会更多地考虑信息的过量存储一样，依赖用户信任的信息服务商也会不得不调整他们的策略。

对于一些公司，像Google，对于存储期限的改变需要一些技术工作，可能甚至会更改商业运作。对于其他的公司，这无非是顺应了他们现有的隐私承诺。以微软为例，它在堪称典范的《软件产品与服务隐私指南》中提到，“用户数据应该被按照最少时间需求而被保留”，“任何在公司存储的用户数据应该有一个保存政策，指明这些数据应该被保存多长时间以及，在何时应该被从所有数据库中清除”。

设定存储期限的技术措施

至此为止，我已经描述了文件的存储期限是如何工作的。但在某些情况下，渐进而细微的方法是更可取的。cookie是一个再合适不过的例子。所有被送到浏览器的cookie被存储在用户计算机中一个精心设计的cookie文件中。但是每个cookie的价值有效期各不相同，因此理应为每个cookie分别设定一个存储期限。有趣的是，对cookie元信息的标准描述已经预见到了cookie存储期限的的来临。我们仍然需要做的只是让用户在收到cookie以后做一个简单的操作来设定时间，或者提供一个在以后更新时间的机会。

亚马逊提供了另一个直观的例子。它的书籍推荐系统基于前一本书的销量和查阅情况。亚马逊已经为顾客提供了一种从用于推荐的信息库，这个库能避免过早推荐信息的影响。但是用户必须定位和浏览特制的网页才能这样做。如果在结账的时候，顾客可以获得一个输入交易信息保存期限的机会，事情就会变得更容易。

打个比方，如果他们要为某人预定生日礼物，他们会选择短保存期限的信息，所以，未来提供给他们的推荐就不会受到这次他们为别人买东西造成的信息“污染”。如果某人正在紧急订购旅行或考试用书，他可以选择一个出行时间或考试时间作为合适的信息期限，这样做不仅节省了存储空间，也优化了未来的推荐的质量。

要求顾客在购买的时候指定一个期限（当信息被添加到用户的文件中）为现有的系统提供了两个优势。第一，一次性为他人购买可以快速地被推荐数据库过滤掉。第二，顾客更喜欢能够对他们何时对何种话题或作者感兴趣有一种经验性的把握，而不需要在几个月后重新回顾这个问题。

像Google、微软、Yahoo！或者Ask.com这样的搜索网站可以同样为网页和用户检索提供一个期限。让我们先讨论网页。目前的网页搜索引擎在互联网上“抓取”信息。他们通常通过网页之间的链接这样做。任何用这种方式新发现的网页会被添加到搜索引擎主页，而且会保存一个副本，以免原件失效（Google称之为“caching”，即“缓冲”）。用这种方法，以巨大的计算和存储能力作为依托，网页搜索引擎为网上有价值的信息建立了一个巨大的索引。在这种网络不断改变的状况下，搜索引擎必须不断地从网上抓信息，并保持其信息不断更新。简而言之，网页包含了浏览器上显示的内容，以及关于如何在视觉上呈现这些信息的指令。因为用搜索引擎定位内容非常重要，网页通常包含为搜索引擎提供的元信息，提供指示其内容的关键词。其中的一种元信息标签可以被用作要求搜索引擎不要引用和储藏特定的网页，因此，能够将它隔绝于数字化记忆。这已经使遗忘变得可行，尽管是一种非此即彼的二元化的形

式。

除了冷酷的非此即彼之外，我们预期会出现由网页作者定义的新的元信息标签。这些元信息标签能够指定搜索引擎可以在未来多长时间内在保留这些网页链接的。因为这些为搜索引擎翻译元信息在技术上的基础构架已经就位，所以增加新的元信息标签就会相对容易和便于使用。结果就是，网页作者可以选择他们的信息被保留的时间。此外，网页搜索引擎甚至可以紧挨着搜索结果显示其期限，这虽然看似粗糙，但实用地为用户提示搜索结果的相对价值。

期限也可以用于查询。在输入一个查询时，用户会被提示输入一个期限，或者从显示的候选项中选择一个。当用户搜索其他人的计算机，或者为同事或朋友提供一次性的搜索，他们就设定较短的期限，比如当这些行为会使预测长期兴趣的算法产生错误结论的时候。用户将不得不花费一些时间去操作，决定或点击——但是作为回报，他们会对他们的搜索历史拥有巨大的掌控力（和更多未来搜索的相关结果），并且减少了数字化记忆。

网页搜索站点，会根据搜索条目存储期限，并据此处理信息，包括自动删除到期的记录。从数量上来说，虽然搜索站点会丢失一些他们曾精心调整的搜索结果排名，但是会因为古旧条目不会溢出进而提升质量。期限的潜在应用甚至超越了个人计算机或网络服务，而包括了紧急应对科技。例如，做紧欧盟讨论了在我们的数字化记忆中使用RFID（调频识别）芯片和类似地网络感应设备。欧盟委员会认为，应该促进“解除RFID芯片或类似地手段，增强了用户的控制”。除了委员会所建议的这种相对严酷的“芯片噤声”之外，或许批准信息存储期限是一种更有益的方法。

正如这些例子指出的，为信息引入一个期限不需要个人用户学习复杂的新用户界面。但设计这种易于使用的界面既是一种挑战，也是一种必须。如果想让期限能被轻松理解，用户必须能够很简单地设定它们——让一个快速的想法和一到两个额外的鼠标点击成为全部。使用何种水平（文件或小的像个人cookie一样的信息形式）的存储期限将取决于信息的类型和使用的环境。

不是用技术删除，而是让遗忘复活

在第4章，我提出了数字化记忆的潜在危险，用了“信息权力”和“时间”这两个词。我在第5章详细描述的一系列针对数字化记忆的主张，都是以保护信息隐私的机制为基础的。我已经指明，遗忘的消亡与信息隐私的原则并不一致。对于让遗忘复活，我的建议可能是通过存储期限，

瞄准阻断数字化记忆的“时间”挑战。信息存储期限的目标是减少数字化记忆中的信息。在这个范围内限制了能被分享给别人的信息，存储期限会有一种信息隐私的净优势。但我必须强调，当我们通常称为信息隐私、机密或知识产权的信息在两个或更多团体之间分享时，存储期限并不能解决在这些情景下的更普遍的对信息的控制。遗忘的消亡与其对信息隐私的威胁足以要求我们做出足够的应对。我认为，存储期限差不多可以帮助缓和数字化记忆带来的挑战，但它永远不能代替信息隐私权的防护机制。

我之前举的例子描述了信息存储期限在不同的场景和情境中的实现。它们的共同之处在于，它们唤起了对信息时间维度的觉察，在将信息存入数字化记忆之前设定一个期限，并自动删除过期的信息。但在这种核心结构之外，我们可以想象，由于类型和情境的不同，信息的存储期限有许多不同的版本。同样重要的是，存储期限的具体实施也必须反映出不同的社会偏好。这些偏好可以从法规结构和我们选择的技术构架中找到。在不同的社会文化中，这种偏好有所不同，就像信息隐私和知识产权立法和实施的程度在不同国家是不一样的。存储期限的相关法律也会由于这些不同而产生不同。

我认为，社会偏好今后至少会通过两种维度表达。基于这两种维度，构建存储期限的不同手段，取决于在社会中带有存储期限的信息是如何被共享的。第一个维度是社会对存储期限的持续性和“黏性”（关联性）的期望，第二个维度反映了社会将会信任谁来设定存储期限。

我在前面曾提到，存储期限是与其指涉的信息相关联的元信息。在那些只想要针对数字化记忆获得最少保护的社会中，立法者只会简单地要求存储期限必须与信息连结。在这样的系统中，如果不受到合约或其他规则的限制，信息的接收者可以任意更改存储期限。技术上讲，这样的应用对信息处理的限制是极有限的。从意识的角度讲，即使在这样弱小的系统中，社会规范和市场压力，即影响人类行为的另一种机制，可能会对人的行为有潜在的影响。假设用户先挑选了存储器，然后将信息分享给其他团体。如果这些信息接收者公开无视自己顾客的喜好，比如延长一些过期信息的寿命，顾客就会把他们蹬掉，把业务交给其他人。当然，这取决于顾客改变供应商的成本，也就是真有其他供应商存在。

如果社会想要在这个最小值的基础上更进一步，立法者应该命令存储期限普遍不可更改，这相当于，它们只可以被创作者或在经其授权的情况下更改。这种规则很好理解，它和隐私权类似。即使是在这种相对严格的系统中，信用管理和直销数据库卖家将继续运作，但这会迫使他们更多地思考信息收集和保留的政策。同样，这会迫使旅游业计算机订

票系统的操作员删除被取消的订票记录，而不是像现在这样长时间地保留它们。此外，政府机构将不得不决定在某时刻删除市民的信息，而不是把所有东西存起来。

在某些情况下个人不适合自己选择存储期限。例如，当个体对设定太长或太短的存储期限的潜在的危险性后果没有觉察时。在这种情况下，立法者会设定一个存储期限的上限，或要求修改某些类型的或某些交易信息的存储期限。

与工程师们讨论这些维护存储期限稳固性的立法选项时，特别是在美国本土之外讨论时，我经常听到一些怀疑这些立法有效性的声音。他们认为，这样的法律不足以保证“粘性”，并指出这和信息隐私权保护法的低效率是类似地。它可以做得更好吗？一些人争辩道，运用技术区确保存储期限的稳固性？除了禁止信息处理商在未经个人同意的情况下修改存储期限，立法者可以要求使用DRM去确保存储期限的稳固性。包含存储期限的信息可以被数字化的标记，使未授权的修改存储期限变得更困难，比如经过第三方。当然，为了避免DRM的漏网之鱼，立法者可以要求禁止未授权的改动或干预存储期限，类似于备受争议的千年数字版权法案（*Digital Millennium Copyright Act*）。

也许，在技术上保证存储期限的稳固性对于特殊的敏感类型（或情境中）的信息是合理的，但是我在上一章已经列举了大量的原因，分析DRM构架，阐述了我对消除人为因素并追求完美方法的保留和怀疑态度。总之，信息存储期限的核心思想，不是通过委以技术手段将数字化记忆的问题推离我们的意识，恰恰相反，让人们觉察到遗忘的价值和重要性。

谁来掌控存储期限

到此为止，我已经解释了用户在数字化记忆中保存信息时存储期限是如何工作的。在许多例子中，信息并不由一个人生成。以电子商务为例，产生的信息同时与卖方和买方关联。那么应该由谁设定存储期限呢？在本质上，这重复了信息控制的争论，不同的只是谁控制信息的存储期限，而不是控制信息本身。如何决定联合生成的信息问题看上去似乎和我刚才列举的持久性问题相似。在两个例子中，问题都是谁来控制，但是其中有一个本质区别。持久性指的是在存储期限被设定后被保存的时间长度。当信息是被共同创造的时候，还没有设定一个存储期限，双方对其设定都有几乎相同的合法性。

那么哪一方设定呢？答案是双方。原则上，每个交易伙伴可以在添加数字化记忆之前，分别决定信息的存储期限，一个人不会因为另一个

人（交易伙伴）的设定而受到限制。表面上，这会有所改进。如果买家和卖家都可以无视交易信息的存储期限设定，Google或亚马逊或任何这样的信息处理商都不必更改操作，遗忘将继续成为偶然，而非原则。即便是这样软弱的存储期限，在实施起来也会产生净收益（虽然很小），出现类似于我所描述的“粘性”。这使得用户面对它们特定的存储期限要求时，觉察度将会上升。也许个人会开始期望存储期限会被他们的交易伙伴接受和评估。在这种程度上，顾客个人会希望并能够将这些期望体现在市场的行情中——对业务伙伴的选择基于他们对顾客的存储期限尊敬与否，信息处理商们挽留客户的欲望将会驱使他们调整其策略。

还有一个更令人信服的方法是，应该为共同设定的存储期限颁布一个正式的法规，以使所有参与者都必须使用信息的相同的存储期限。因此个人可以知道他们共同创建的存储期限就像图书购买记录或网页搜索记录。这种透明度以及对数字化记忆潜在危险的觉察，可以协助个人展开针对贸易伙伴关联的存储期限的谈判。在亚马逊图书销售记录中，亚马逊可以为客户提供一个日期排序，因此顾客可以挑选一个合适的日期。当然，围绕针对存储期限的谈判通常是贸易谈判的一部分，如果交易伙伴不能就共同的存储期限达成一致，就不会发生交易。这也会损害卖家，而不仅仅是顾客，特别是在经济萧条的时期。

但是，把希望完全寄托在市场压力上似乎不太合理。联合设定存储期限的方案预先假定了交易各方有大致相同的谈判筹码。然而，这就像我在第4章提到的，这是一个不太现实的情境。当参与到存储期限的谈判时，个人可能经常处于被动，并且他们可能也只有有限的申诉的渠道。这是在信息市场上十分普遍的现象：当有数十或上百的卖家时，只有一小部分卖家占领了大部分市场。尤其是建立在互联网之上的公司，成功会导致更大的成功（基于网络效应），因此极大地促进了强势卖家的形成。在另一方面，顾客变更的成本比在一些传统情况下更低，可能抵消了一些不平等谈判的影响。剩下的不平衡可能会促使立法者提供一个更渐进的针对共同创建存储期限的解决方案。特别是在谈判信息权力悬殊的情况肯定会被注意到。有一系列的选择，比如根据情况，法律可以设定一个允许的存储期限范围或上限，或者在特定情况下的特定交易时，服务商必须接受顾客设定的存储期限。

共同存储期限的谈判可能会由于技术构架而变得便利。比如，通过使用类似于爱德华·费尔顿（Ed Felten）和他的同事提出的已经极简单的系统（我在第5章描述的）。对于用户，这种谈判不可避免地涉及到复杂的协议。不妨思考一下拍照和摄影的情况。

假设某人给你拍了一张数字照片，如果这被看作一个信息交易，这

张照片的存储期限应该是在你和拍摄者之间共同设定的。这种谈判应该由机械地设定，而不用麻烦人。每部数码相机可以嵌入一个选择存储期限的程序，可以通过一个简单的设定实现。在拍照之前，照相机发送一个“拍照请求”。想象一下，假如我们携带小的“授权装置”（钥匙环大小），当接收到“拍照请求”时回复一个预设的存储期限。这个装置甚至具备一个小的选择器，亦可快速更改预设的存储期限，可能从“零”（如果某人不想被拍照）到“从现在往后的三年”（一个中等时程）再到“100年”（一个值得纪念的事件）。

当照片被拍摄时，它将按照接收到的最短的存储期限存储（包括拍照者设定的）。照相机甚至可以提供一个可视信号，指示是否有谁的设备说了“零”或者类似与参与者有关的极端值。这种解决方案可以被大多数用户实施，使得不想被长时保存的照片会被逐渐遗忘。

显而易见的是，存储期限并不是个一蹴而就的方案。可以想象，一整套选项要从把核心观念一步步落实，然后逐渐形成复杂的、有些长远考虑的版本。各种实施方案都有很广泛的变化，没有一种版本是天然优于其他的。社会选择哪个版本不仅取决于信息的内容和上下文，也取决于社会的期望和价值观，以及负责推广它的机构。

我们需要“能衰退”的存储系统

和其他替代品相比，信息的存储期限有一些优势。和“数字化节制”不一样，它们参与到了数字文化和全球网络当中。与信息隐私权相反，存储期限不会单单依靠个人在时间和金钱上的消耗来在法庭抗争。存储期限也不需要与DRM系统一样的完全是技术应用的控制。与数字化节制、信息隐私权与**DRM**系统相比，存储期限的一个优势是，它解决了信息隐私的问题，改变了与数字化记忆相抗争的意图。存储期限的目的用来对付数字化记忆的“时间”维度，不会对不想处理的信息存储数量进行限制。存储期限的确需要被广泛地意识到，但不是一种认知革命。它比信息生态的法令更宽泛，但比一个完全视情况而定的构架更局限。

存储期限是对遗忘消亡的温和回应，至少通过四种方式。首先，从技术上讲，它们使用了许多不同的已经存在的理论、实践和机制，在许多例子中，只需改变有限的技术。其次，从法律角度讲，它们引入了数字领域默认的遗忘，这对我们来说是很熟悉的，而不依靠新的权力和机构。第三，存储期限在规范人类行为方面也是温和适度的，包括法律和软件。第四，政治上讲，它们比广泛的管制方法更容易接受，比侵略性的DRM系统受到更少的争议。

虽然在一个重要的方面，存储期限看上去似乎是激进的：它们以改变数字时代记忆的默认性为目标。因为遗忘是人类决策的核心，这种改变会对数字生活的塑造产生极大的后果。

通过存储期限重新引入记忆也有一系列真实的、可觉察的短处。

第一，一些人可能不喜欢存储期限，因为他们没有任何信息隐私权的诉求。这是事实。但像我提到的，他们不应如此。

第二，存储期限会因为妨碍工作而被指责，进而导致对档案馆和图书馆的存在的质疑。这种想法是无根据的。存储期限让个体能够决定他们想要记住的一些信息的时间长短。但社会有权利在必要的时刻无视这种个人选择。例如，通过授权保留信息，通过保留图书馆、档案馆和其他特别机构去保存一些特别重要的关于过去的事件的信息。这个提议不会改变如下原则——除了保存信息的需求十分明确和明显时，才存在例外。

第三，更普遍的是，个人的遗忘欲求和社会的记录欲求将会发生冲突。在个人主义比集体主义更被推崇的社会中，正如原始的经验法则一样，个体应该是设定存储期限的一方。然而，在某些情况下，社会考量会比个体期望更重要。例如，最近欧洲的法院不得不做出选择，要么公布过去共产主义告密者的姓名，要么在冷战结束二十年后让他们受益于遗忘。这是不能用存储期限替代的决定；这是一个社会必须做出的决定。如果一个社会希望阻止特定的历史事件的遗忘，它可以授权（长时程）存储期限，或者为了尊重存储期限而豁免一些信息档案馆。只要这种授权属于例外情况，而非普遍规则，它们就与存储期限的原则一致，遗忘仍然是默认的。

第四，信息的存储期限不能保证完美。技术锁死将使逃避存储期限变得不可能，如果缺乏这种技术锁死，就会存在不同程度的漏网之鱼——信息和存储期限取消关联以避免被删除。但正如法学家们一直坚持地那样，完美的强制力不是广泛服从的必要条件。有一些强制力已经足够，特别是对于被社会广泛接受的规则。如果对逃避存储期限会被社会所厌恶，又动用了诉诸法律的强制和高压的技术手段，遵循存储期限的效果就会很好。

第五，正如我之前提到的，设定存储期限不是为了动用信息权力消灭异端。例如，如果存储期限是被两方或多方共同设定的，信息权力更大的一方会推动信息权力更小的一方同意存储期限。立法者需要审慎地制定一套支持性的法律框架用于对抗这种后果，或至少减少频度和严酷性。我之前构想了这样的法律框架的一些可能的元素。但即使是支持性的法律框架，一些不平等的情况仍会持续。毋庸置疑，这是存储期现的

缺陷所在。然而，这种缺陷并不严重。存储期限的主要目标是针对数字化记忆的“时间”挑战，而不是“信息权力”挑战。其他针对信息权力的方法普遍会为存储期限增加更平等的谈判地位。存储期限可能会结合一些我在第5章提到的应对方式，特别是在那些对抗数字化记忆在“信息权力”方面的挑战。同样重要地是，存储期限减少了别人存储的关于我们的信息的总量，因此限制了别人优于我们的信息力量。这反过来会使存储期限的谈判更加平等。

第六，许多存储期限的成功将取决于用户设定的容易程度和直观程度。只有当它们不被看成是烦人的事物，而是有价值的提醒物时，它们才会被接受。

第七，在大多数简单形式下，存储期限受制于结构性的弱点。它们天生就是非此即彼的：信息在存储期限之前是可获得的，而在之后完全是无法获取的。其中没有一个渐变，没有一个在完全遗忘其核心内容之前，失去细节的过程。在这种意义上，存储期限是和人脑中的遗忘机制不同的。

在某种程度上不可避免地是，我们对人脑的遗忘和回忆机制了解得还不够多，不足以精确地用数字代码复制它们。但这不能阻止我们去尝试。也许在只增加少量复杂性的前提下，是能够构建一个更接近人类记忆和遗忘功能的存储期限机制。例如，我们可以使我们的数字存储设备察觉到我们调用数据的频率，并在设备中调整相应条目的数字寿命。通过这种方法，可以模拟人类记忆。同样，信息处理器可以为用户提供视觉线索，指明一组数据接近存储期限的程度，以使用户更容易应对。

取舍之道

从长远看，更值得期待的版本应该是“能衰退”的记忆。我们可以想象，旧信息需要花更长的时间才能被搜索出来，就像我们的大脑有时需要额外的时间来回忆起久远的事件。或者数字化记忆会需要更多的查询信息才能找到更久远的信息，这模仿了我们的大脑有时需要更多的刺激才能回忆的情况。一个文件也可以随时间变得模糊或被淡化，而不是突然被删。

我们非常有必要继续进行理论性地尝试，去打造精致和复杂的数字遗忘系统。与那些遗忘的完善方法相比，存储期限是粗糙的。但在缺乏更完善的方法的今天，存储期限将向一个更善于遗忘的世界迈出有价值的第一步。

这一章集中讨论了数字时代对遗忘的重新引入。这是目前一个可能的应对数字化记忆潜在威胁的方法，它用于应对我特别在第4章提到的“时间”的挑战。现在存在许多可能的建立数字遗忘的方案。其中一个

就是通过设置存储期限重新引入遗忘，以及遗忘的默认性。存储期限对于不同情境中的许多类信息都比较有效，并比其他方法具有优势。它以提升人类觉察力和权力为目标，需要相对少的技术挑战。它通过清除过时和无关的信息，提升了数据存储的整体质量。存储期限能够为不同的社会预期和价值观提供弹性。从长期看，存储期限会带来一种缩小信息权力差距的净收益。更重要的是，存储期限逐渐为我们担负起了理解数字化记忆的重担，并为人类和社会保留了“快速反应”的能力。

然而，存储期限确实存在一些弱点。它并不完美，不能消除或增加个体遗忘期望和社会记忆期望的矛盾。在需要共同决定的范围内，信息权力的差异不可克服。这种情况似乎需要附加的手段（包括立法）来应对。存储期限若要长时间地成功，极其依赖易于操作并高效的用户界面。最后，存储期限天生就是相对二元化的——当达到一个期限，信息就会被删除。在未来，我们可能会想要用一个更复杂和渐变的遗忘机制来替换它，允许信息随着时间逐渐“分解”或“锈蚀”。从一般意义上讲，有必要理解存储期限不是万能的。我建议，不要取消所有恰的针对数字化记忆的措施，而是包容它们，建立一种强大的组合。

数字时代能否遗忘，特别是存储期限能否实施，在很大程度上取决于我们的选择。社会，以及我们每个人是否愿意迈出这必须的一步，将需要持续的讨论，以决定建立公众认同的具体机制。它需要在各个方面做出的努力，就像知识产权法律的改革那样。也许这些努力的最初几步已经迈出了。比如，在阿根廷，作家亚历杭德罗·托莱多里尼（Alejandro Tortolini）和他的同事恩里克·科吉兰诺（Enrique Quagliano）已经发起了一场名为“重新发明互联网遗忘”的运动。他们已经通过电视、广播和平面媒体上呼吁公众，持续地完善了他们的思想，并意外地获得了有力的公众支持。

在这个振奋人心的时刻，在全世界的许多社会中，仍然有很多很多事情要做，才能增加社会支持，并使存储期限与其他使遗忘复活的手段相辅相成。这一切将使数字时代更加人性化。

07 让遗忘回归常态：大数据时代数字化记忆的未来

人类必须走出数字化记忆的阴影，而我们也有能力做到，那就是恢复遗忘。这并不是在鼓吹一个无知的未来，而是在承认一个随着时间推移思维会演变、观点会调整的未来。无论如何，面对即将到来的大数据时代，我们要确保自己依然记得遗忘的美德……

遗忘在人类决策中起到了重要的作用，它允许我们从个人经验中去概括和抽象。它使得我们能够接受自己，就像所有的生命一样，会随着时间推移发生改变。

遗忘让我们瞄准当下，而不是将我们永久地拴在一个越来越无关的过去里。

作为人类，我们不能无知地穿越时间去旅行。因为我们有记忆的能力，所以我们能够去比较，去学习，去体验时间的变化。同样重要的是，我们有遗忘的能力，有将我们自身从过去桎梏中解放出来的能力，有活在当下的能力。过去几千年，记忆和遗忘之间的关系一直很清楚。记住很难而且代价高昂，因此人类不得不谨慎地选择哪些值得记住。人类默认的状态是遗忘。在数字时代，或许人类发生的最根本的改变，就是记忆与遗忘原有平衡的反转。将信息提交给数字存储器已经成为默认状态，而遗忘则成了例外。

数字化使得存储成本的垂直下降、简便的信息提取，以及全球性访问数字化记忆成为可能。在人类历史上，这是第一次我们能够让记住比遗忘更便宜也更容易，也是第一次逆转了遗忘由来已久的默认状态。原因显而易见——即使人类自身的记忆经常失效，难道我们大家都不想选择拥有全面的记忆吗？难道我们大家都不想给子孙后代留下我们的思想和印迹吗？如今，我们发现已经身处一个拥有完整数字化记忆的“美丽新世界”。

在这个世界中，诸如Google之类的信息处理器让全世界能够访问海量信息，从照片到博客，从详细的市场信息到我们自家（和邻居家）后院的高清卫星图像。很快，类似史黛西·施奈德和安德鲁·费尔德玛的案例将会成为普遍现象：过去微不足道的行为被完全记忆下来，生活因此被打乱（或者至少是被干扰了），个人被暴露在苛刻到不可思议的社会面前，这不仅仅发生在北美，而是在全世界。事实上，在这本书的写作过程中，德国一个小小的保守派政治家就被解雇了，因为在一个类似facebook的社交网站上显示，过去他曾加入了一个名字和内容都有些可疑的在线团体。

在第3章中，针对为什么这个趋势将与我们有关，我解释了两个重要原因。首先，信息隐私专家指出，它进一步将权力从被监视者转移到监督者身上，同时跨越时间的维度。的确如此。数字存储器的发展，让我们能够相对容易地重塑言行，即便它们已经过去了很久，它们建造起来的不仅是空间上与时间上的"圆形监狱"，而且还约束我们表达想法和参与社会的意愿。我们真的想生活在一个充满了奴性和恐惧的社会中吗？

其次，遗忘在人类决策中起到了重要的作用。它允许我们从个人经验中去概括和抽象。它使得我们能够接受自身，就像所有的生命一样，会随着时间推移发生改变。因此，它让我们瞄准当下，而不是将我们永久地拴在一个与现实越来越无关的过去里。而且，遗忘能使社会原谅它的成员，并且对改变保持开放的态度。数字化记忆渐渐破坏了遗忘所起到的重要作用，从而同时在个人和社会两个层面上威胁到我们学习、诠释和适时行动的能力。它还让我们暴露在人类具有潜在毁灭性的过度反应面前——那就是另一个极端，完全漠视我们的过去。

我们必须应对数字化记忆造成的挑战，而且我相信我们能够应对，能够恢复遗忘的能力。我并不是在鼓吹一个无知的未来，而是一个承认随着时间的推移人类会改变，想法会演变、观点会调整的未来。我们能够采用多种不同的方法，应对数字化记忆的阴影。在第5章中，我概述了其中的六种方法——共两组，每组三种。在每一组中，一种应对方法都通过社会准则、法律、技术架构去培养遗忘。第一组应对方法（衍生自信息隐私辩论）主要解决了我称之为数字化记忆在"信息权力"上的挑战，而第二组应对方法则瞄准了数字化记忆在"时间"上的挑战，因此可以说更为直接地面对了遗忘的终止。虽然所有六种策略每种都许下了一些承诺，并对我们面临问题的复杂性提供了深刻的见解，但是没有一种策略是"万全之策"。为了有效地应对数字化记忆，我们可能不得不将这几种方法联合起来，或者甚至加入进一步的应对策略。

在第6章中，我又提出了另一种应对方法：在数字时代通过给信息设置有效期限，重新引入遗忘的概念。目的是为了将默认状态从永久保存信息调回到可以在一段时间之后删除它。我描述了存储期限在各种结构上、法律上、技术上的组成部分，以及它们如何协同作用。我提供了一系列可能的执行措施，这些措施都是基于政策制定者和公众想彻底恢复遗忘的渴望程度。存储期限在若干措施构成的维度上是相对温和的一种，使得这种方案相较而言更容易被采纳。然而，它可能已经足够阻止并逆转这种转向记忆的趋势，同时恢复我们遗忘的能力，这对于人类来说至关重要。虽然我相信通过有效期限能促进遗忘的有效性，但我也

承认这一方案同时会伴有与生俱来的缺点，而且无法解决记忆的所有问题。无论如何，在现实中，它可能会被证明确实是我们所需要的，是现有的一系列应对措施中有益的补充。

不过，最为重要的是，我想让大家开启一场关于“遗忘”的广泛、开放、热烈的讨论，同时这也关系到怎样才能确保人类在进入大数据时代时，还会记得遗忘的重要性，记得遗忘是一种美德。

后记

所有这一切都开始于2007年2月，那半页名为“被遗忘的权利”的笔记。虽然后来我很快就忘掉了那半页笔记，但却记住了这个思想。当我的记者朋友肯尼思·库克耶（Kenneth Cukier）打电话问我关于无处不在的计算机技术与隐私的新视点时，我向他讲述了关于遗忘的重要性。之后，他将这个观点写进了他的报道中。

2007年4月，我的研究重点迅速被媒体获知，这多少有些出乎我的意料。我的想法似乎切中肯綮。在这两年中，我逐步形成了我的论证，我的观点已被传播开来。而且让我高兴的是，有很多人要么认同这个观点，要么正在提倡这个观点。

在隐私专家和从业者的圈子里，有很多人为我提供了非常有价值的信息。我尤其要感谢杰弗里·弗莱德伯格（Jeffrey Friedberg）和保罗·施瓦兹（Paul Schwartz）。艾琳娜·萨普莱基娜（Elena Saprykina）在俄罗斯历史方面为我提供了很多帮助。马克·罗滕伯格（Marc Rotenberg），尼古拉斯·伦布拉德（Nicklas Lundblad）和菲利普·缪勒（Philipp Müller）为本书提出了简洁而有价值的评论。爱德华·费尔顿（Edward Felten）、让-弗朗索瓦·布兰切特（Jean-François Blanchette）和利亚姆·班农（Liam Bannon），他们都在遗忘和记忆方面有着极富见地的观点，他们都为我提供了很有帮助的指引与最受欢迎的见解。我在新加坡国立大学李光耀公共政策学院信息与创新政策研究中心的同事特蕾西·劳（Tracy Loh）阅读了整部书稿，并对全书提出了许多有价值的建议。我要特别感谢罗恩·伯特（Ron Burt）与迈克尔·加里（Michael Curry），他们花了大量时间去阅读我的论证，他们以极大的鼓励和发人深省的见解对我的论点进行了全面的评论。此外，我还要感谢三位匿名评审人极有帮助的建议。

我很幸运能够将我的论点不断推敲并呈现给全世界的读者。特别感谢让·弗朗索瓦·布兰切特和他在加州大学洛杉矶分校的同事们，汤姆·布鲁斯（Tom Bruce）及其在康奈尔大学的同事们，托马斯·霍伦（Thomas Hoeren）及其在德国明斯特信息、电信与媒体研究中心的职员，桃瑞斯·欧贝拉德（Doris Obereder）及其在奥地利电子艺术节工作的同事，以及马库斯·贝克达尔（Markus Beckedahl）。同样重要的是，我极大地受益于最近两年来我与全世界传统媒体和新媒体记者们的多次交谈；他们的问题帮助我塑造并发展了我的论证。尤其是科技博客网站Ars Technica的纳特·安德森（Nate Anderson），《华尔街日报》的李·戈麦斯（Lee Gomes），《纽约时报》的克拉克·霍伊特（Clark Hoyt），以

及德国《时代周刊》（*Die Zeit*）的凯·比尔曼（Kai Biermann），他们所有人都帮助我看到了之前我没有意识到的那半页笔记的要旨。我还受到了亚历杭德罗·托莱多里尼（Alejandro Tortolini）和恩里克·科吉兰诺（Enrique Quagliano），以及比尔·施特劳斯（Bill Straus）的鼓励，前两位在阿根廷发起了一场重拾遗忘的运动，施特劳斯是美国马萨诸塞州推动类似运动的立法者。

如果本书有一个最核心的推动者的话，那么这个人非蒂姆·苏利文（Tim Sullivan）莫属。蒂姆是一位极好的编辑，是他第一次让我想到将那半页笔记扩展成一本书。即便在他从普林斯顿大学出版社离职后，他仍然继续密切关注我的进展。在蒂姆离职后，查克·迈尔斯（Chuck Myers）以极大的热情继续着这个项目，我非常感谢他中途接手这本书并一直坚持到它出版。

现在我已经忘记了写作过程中的许多痛苦，这对一本讲述遗忘美德的书是非常合宜的。我将本书献给波吉特·拉比尔（Birgit Rabl），她不仅是本书的第一位读者与批评者，也是我生命中的挚爱。